DNA Methylation in Rectal Cancer: A Genome Wide Study

Rachel Dbeis

Submitted by Rachel Dbeis to the University of Exeter as a thesis for the degree of Masters by Research in Medical Studies (MbyRes), July 2017

I certify that all material in this thesis which is not my own work has been identified and that no material has previously been submitted and approved for the award of a degree by this or any other University.

Signature: 

Abstract

Rectal cancer differs from colon cancer in terms of its underlying biological behaviour, clinical course, genetic and epigenetic aetiology. Epigenetic mechanisms modify gene expression independently of DNA sequence. DNA methylation is the most studied epigenetic mechanism, known to play a role in colorectal cancer. The role of DNA methylation specific to rectal cancer however, is poorly understood. Here we present the results from a DNA methylation study of 45 individuals with rectal cancer.

A total of 45 patients (>18 years) diagnosed with rectal adenocarcinoma (stages II and III) who underwent or were undergoing treatment were recruited. Matched rectal tumour and adjacent normal mucosal samples (n=90) were obtained from each patient and processed fresh frozen (n=32) or embedded in paraffin (n=58). DNA was extracted and checked for quality and quantity, treated with sodium bisulfite and run on the Illumina Infinium HumanMethylation 450 Beadchip. Only samples that passed the Quality Control were subsequently analysed (n=30).

A combined linear regression analysis of all 408,652 probes showed that global levels of DNA methylation are decreased in rectal cancer samples compared with normal unaffected samples. In total, 176 differentially methylated probes and 828 differentially methylated regions were identified in rectal cancer vs normal tissue. All the genes identified underwent gene ontology analysis to assess whether they are biologically meaningful.

In summary, our study focused on the discovery of *de novo* epigenetic changes associated with rectal cancer, using a genome wide approach and novel bioinformatics approaches. These findings improve our understanding of the epigenetics of this disease. Furthermore, they have the potential to be used as biomarkers for detection, prognosis and monitoring treatment response in patients with rectal cancer.

The work presented in this thesis was funded by the Jean Shanks Foundation, the Royal College of Surgeons and Bowel Cancer West.

Approximate word count: 38,000

Acknowledgements

First and foremost, I would like to thank my supervisors, Professor Jonathan Mill, Mr Neil Smart and Mr Ian Daniels, for their extraordinary supervision, support and encouragement throughout my MbyRes. Jon, I owe much gratitude to your support, knowledge and patience. Mr D, your talks have inspired me to aim higher and work harder. Neil, I don't know that I would be here today without you. Thank you for being a mentor, a teacher and a friend.

To my collaborators on this project, specifically Dr Ian Chandler, Bridget Knight and Lidia Romanczuk, thank you for your contributions and hard work. Dr Sarah Saunders, I would like to thank you for saving the day with your energy, efficiency and enthusiasm.

I also wish to thank all current members and alumni of the Complex Disease Epigenetics Group, especially Adam, Emma, Joana, Katie, Helen, Aaron, Becks, Conor, Therese, Isabel and Matt. I specifically would like to mention Eilis for the teaching, patience and support. Your work and help have been central to achieving this project. Beth and Liv, thank you for the science and the laughs. Jojo, king of the lab, you kept me sane. Thank you for helping me but mostly for being my friend. It has been a pleasure and a privilege to be part of such a talented and wonderful team. I also would like to thank the National Institute for Health Research (NIHR), the Exeter Clinical Research Facility, the Exeter Surgical Health Service Research Unit and the kind and helpful faculty at the University of Exeter Medical School, especially Kate Ellacott and Karen Mattick.

I am forever grateful to the Jean Shanks Foundation, the Royal College of Surgeons and Bowel Cancer West who funded and supported this project.

I've been blessed with many people in personal life who have provided me with endless support and love. This page can't contain you all, nor can words express my gratitude. I am thankful for you all.

Katie, you've been incredible throughout. There is far too much to thank you for to begin.

List of Contents

List of Tables

List of Figures

Declarations

All laboratory work was conducted by me at the University of Exeter Medical School laboratories with the exception of the following:

- assistance for the PCR validation of DNA methylation in chapter 2 was given by Bethany Crawford, a PTY student in our laboratory, with further assistance from Dr Emma Dempster.
- Assistance In operating and following the Illumina 450k array laboratory protocols was provided by Dr Joe Burrage.

The bioinformatics and statistical analyses were performed by me with exception of the following:

- assistance for the Quality Control checks in Chapter 2 was given by Olivia Knox, a PTY student in our laboratory, with further assistance from Dr Eilis Hannon.
- The statistical analyses were carried out with help from Dr Eilis Hannon.

List of Abbreviations

Abbreviation	Term
aFAP	Attenuated familial adenomatous polyposis
AJCC	American Joint Committee on Cancer
bp	Base pair
cCR	Complete clinical response
CIMP	CpG island methylator phenotype
CIN	Chromosomal instability
COX2	Cyclooxygenase-2
CRC	Colorectal cancer
CRM	Circumferential resection margin
CRT	Chemoradiotherapy
CT	Computed tomography
CTC	Computed tomography colonography
DFS	Disease Free Survival
DMP	Differentially methylated probe
DMR	Differentially methylated region
DNA	Deoxyribonucleic acid
DNMT	DNA methyltransferase
ECM	Extracellular matrix
EMVI	Extramural vascular invasion
ExeterCRF	Exeter Clinical Research Facility
ETB	Exeter Tissue Bank
FAP	Familial adenomatous polyposis
FF	Fresh frozen
FFPE	Fresh frozen paraffin embedded
FOBT	Faecal occult blood test
GO	Gene ontology
HNPCC	hereditary non-polyposis colorectal cancer
IGF	Insulin growth factor
JPS	Juvenile polyposis syndrome
LARC	Locally advanced rectal cancer
LARS	Low anterior resection syndrome
MAP	MUTYH-associated polyposis
MDS	Multi-dimensional scaling
MDT	Multidisciplinary team
MRF	Mesorectal fascia
MRI	Magnetic resonance imaging

mRNA	Messenger RNA
MSI	Microsatellite instability
ncRNA	Non-coding RNA
nCRT	Neoadjuvant chemoradiotherapy
NHS	National Health Service
NICE	The National Institute for Health and Care Excellence
NIHR	National Institute for Health Research
nRT	Neoadjuvant radiotherapy
NSAIDs	Non-steroidal anti-inflammatory drugs
OS	Overall Survival
PCA	Principle component analysis
pCR	Complete pathological response
PCR	Polymerase chain reaction
PGE2	Prostaglandin E2
PJS	Peutz-Jeghers syndrome
QC	Quality Control
QN	Quantile-normalisation
RDE	Royal Devon and Exeter
RDETB	Royal Devon and Exeter Tissue Bank
RNA	Ribonucleic acid
RT	Radiotherapy
SNP	Single nucleotide polymorphism
TBE	Tris-borate EDTA
TGF-β	Transforming growth factor-β
TME	Total Mesorectal Excision
TNM	Tumour Node Metastasis
TRG	Tumour regression grade
tRNA	Transfer RNA
U.K.	United Kingdom
VEGF	Vascular endothelial growth factor

Chapter 1 - Introduction

1.1 THE COLON

1.1.1 Anatomy:

The large intestine is divided into the caecum, colon, rectum and anal canal. The colon extends superiorly from the caecum and is subdivided into four parts: the ascending, transverse, descending and sigmoid colon (Warwick et al., 1985) (Figure 1.1). The blood supply to the proximal colon arises from the superior mesenteric artery, whereas the distal colon gets its blood supply from the inferior mesenteric artery. The proximal and distal divide is marked by the splenic flexure, situated at the junction of the transverse and descending parts of the colon, where the anastomosis between the two vascular territories occurs (Warwick et al., 1985).

The extrinsic and the intrinsic nerves supply colonic innervation. Extrinsic innervation involves the sympathetic and parasympathetic pathways, which are responsible for colonic motility and sensation. Intrinsic innervation arises from the poorly understood enteric nervous system (Szmulowicz and Hull, 2011). The parasympathetic supply to the proximal colon arises from the vagus nerve whereas fibres from S2-S4 innervate the distal colon and the rectum (Li and Lai, 2009, Szmulowicz and Hull, 2011).

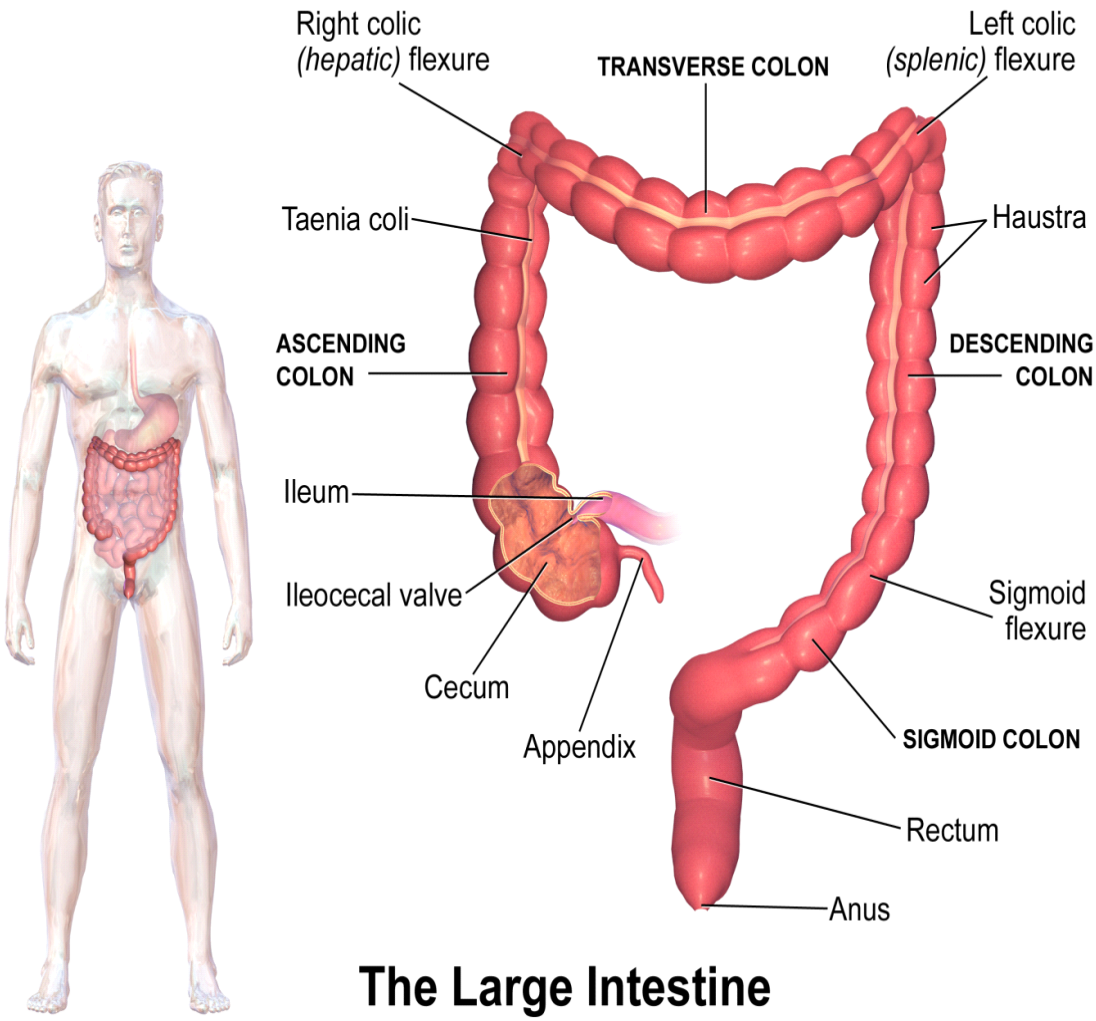


Figure 1.1 The large intestine. This picture shows the different parts of the large intestine and its anatomical subdivisions. Adapted from Blausen (2014).

1.1.2 Physiology and function:

Absorption of nutrients and water along with faecal storing are the main functions of the colon (Li and Lai, 2009). Various complex carbohydrates and to a lesser extent proteins, are processed in the proximal segment of the colon, where nutrients from these products are salvaged via fermentation. The residual products of fermentation and dietary fats that reach the colon are then expelled with the stool (Szmulowicz and Hull, 2011). The enteric nervous system is responsible for colorectal contraction and is under the influence of gut hormones. The varying circulating concentrations of these hormones affect the contractile activity. For example, after meals, the motility of the colon rises

considerably due to cholecystokinin secretion, whereas sleep reduces colonic activity and a high fibre diet helps to retain water and increase the faecal bulk, thus aiding in defecation (Irving and Catchpole, 1992).

The colon is also integral to maintaining appropriate hydration and electrolyte balance through absorption and secretion of water and electrolytes. Sodium and water absorption are greatest in the caecum and decrease progressively towards the rectum. The majority of colonic activity is centred on promoting the salvage of water and electrolytes. Periodically this activity shifts to allow for expulsion of stool through colonic contractions (Szmulowicz and Hull, 2011). Figure 1.2 illustrates some of the physiological functions of the colon.

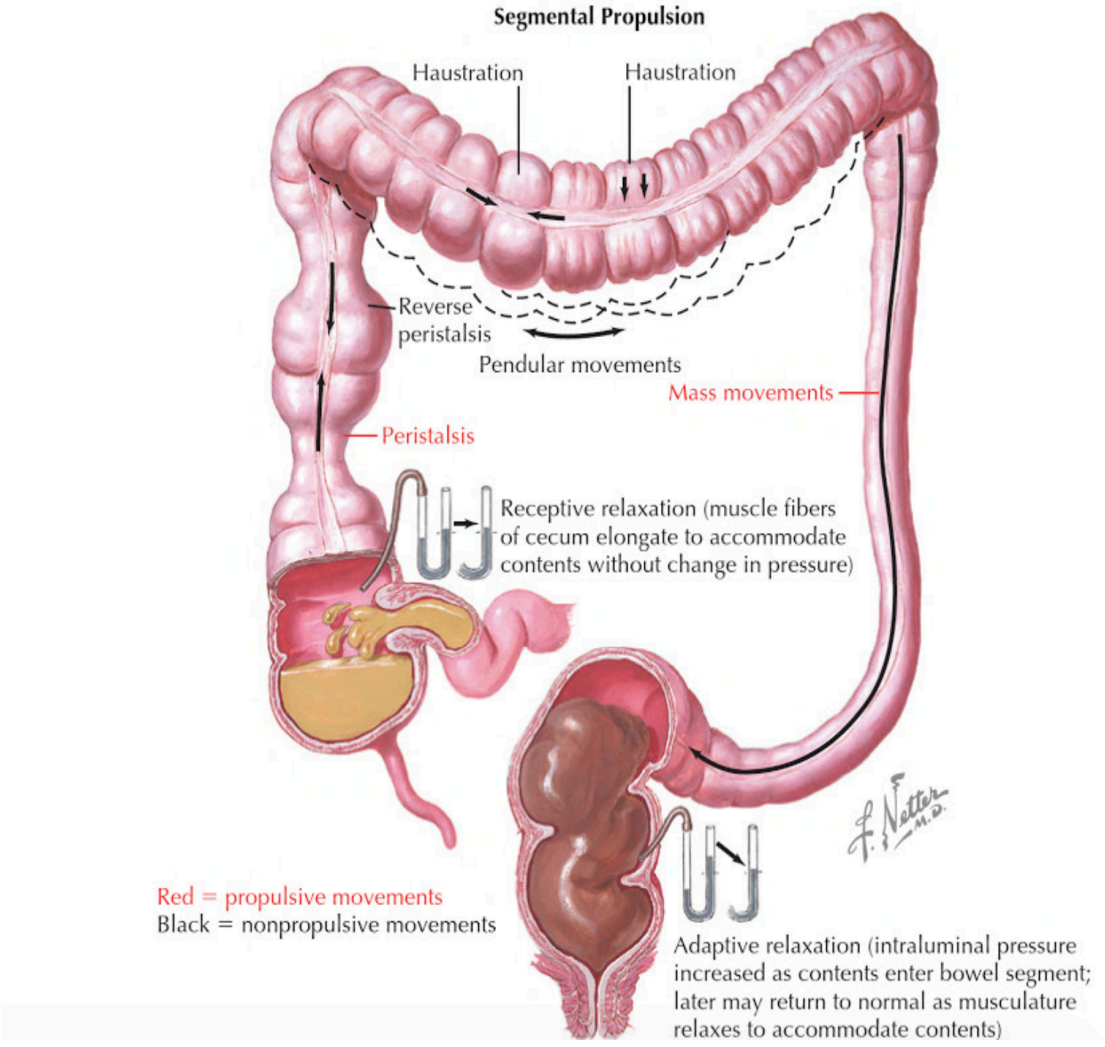


Figure 1.2 Physiological functions of the colon. Adapted from Mulroney and Myers (2015).

1.2 Colorectal Cancer

1.2.1 Epidemiology

Colorectal cancer (CRC) is any cancer that starts in the colon or the rectum. It is the second leading cause of cancer-related death in the USA (Siegel et al., 2017) and one of the most common cancers in the United Kingdom (U.K.) after lung and breast (NICE, 2014). The incidence of the disease is increasing globally, with over one million new cases diagnosed worldwide and approximately 40,000 new cases registered in the U.K. yearly (NICE, 2014, Siegel et al., 2017). By gender worldwide, CRC is the second most common cancer in women (9.2%) and the third in men (10%) (IARC, 2014). The occurrence of CRC is strongly correlated with age, with approximately three quarters of cases occurring in people aged 65 years or older (NICE, 2014). The median age of diagnosis in 70 years in developed countries (Siegel et al., 2017). The prognosis of CRC is slowly improving with improvements in care. 5-year relative survival has reached almost 65% in high-income countries but remains less than 50% in low-income countries. Stage at diagnosis remains the most important prognostic factor (Brenner et al., 2014). There isn't a single risk factor that accounts for most cases of CRC, but several risk factors have been identified such as: family history of CRC, obesity, excessive alcohol consumption, inflammatory bowel disease, smoking, high consumption of red meat and diabetes. Preventative factors include Aspirin, physical activity and hormone replacement therapy (Brenner et al., 2014).

1.2.2 Aetiology

1.2.2.1 Histology

On a cellular level, CRC is thought to originate in multi-potential stem cells found in the intestinal crypts. These stem cells reside at the base of the intestinal glands and give rise to 'transient-amplifying' cells (that become increasingly differentiated with successive rapid divisions, finally giving rise to terminally differentiated cells). At the mouth of the crypts, these cells differentiate into three types: goblet, enteroendocrine and Paneth cells. These differentiated cells eventually undergo apoptosis to the lumen after 3-5 days. Any imbalance in the proliferation, differentiation or apoptosis process within the crypts can lead to aberrant crypt foci, which may progress to adenoma

(Hammoud et al., 2013). The histological progression of these mucosal cells from benign adenomas to malignant carcinomas is characterised by underlying genetic and epigenetic alterations, affecting oncogenes, tumour suppressor genes and DNA repair genes (Coppedè, 2014). The initiation and progression of CRC is therefore considered a multistep tumourigenesis process known as the adenoma-carcinoma sequence (Fearon, 2011).

1.2.2.2 Genetics

Sporadic CRCs, which are likely due to random mutations and devoid of any familial or inherited predisposition, account for approximately 70% of all CRC cases. The second type of CRC is the familial type. In affected families, CRC develops too frequently to be considered sporadic, but does not follow a pattern consistent with an inherited syndrome. The third type is the inherited genetic CRCs, which include: familial adenomatous polyposis (FAP), Gardner Syndrome (a variant of FAP), attenuated FAP (aFAP), Lynch syndrome (hereditary non-polyposis colorectal cancer or HNPCC), MUTYH-associated polyposis (MAP). The rarer ones include hamartomatous polyposis conditions (Peutz-Jeghers syndrome (PJS), juvenile polyposis syndrome (JPS) and others) and hyperplastic polyposis (Migliore et al., 2011, Coppedè, 2014, Nojadeh et al., 2018, Lee et al., 2014). CRC is therefore considered a heterogeneous malignancy with several subtypes, each characterised by different genetic, cytogenetic and epigenetic changes (Migliore, 2011, Perea et al., 2015).

1.2.2.3 Cytogenetics

Genomic instability is considered a key hallmark in CRC. There are three recognised major molecular subtypes: the chromosomal instability (CIN) pathway, the microsatellite instability (MSI) pathway and the recently identified epigenetic CpG island methylator phenotype (CIMP) (Lengauer et al., 1997, Toyota et al., 1999, Migliori, 2011, Perea et al., 2015).

The CIN pathway

CIN is the most common type of genomic instability, occurring in 80%-85% of CRCs (Grady and Carethers, 2008). CIN refers to an accelerated rate of gains or losses – in large portions or in whole chromosomes – that leads to variability in the karyotypes of cells. This causes an imbalance in chromosome number and a high frequency of loss of heterozygosity. The mechanisms underlying CIN include telomere dysfunction, alterations in chromosome segregation and DNA damage response (Marmol et al., 2017). CIN tumours are characterised by mutations in tumour suppressor genes and oncogenes that activate pathways critical for CRC initiation and progression (Lengauer et al., 1997, Migliori, 2011).

The MSI pathway

Microsatellites are stretches of repetitive DNA sequences, which are normally stable. MSI (or replication error positive RER+), the second major pathway, is defined as a relatively frequent change of the length of these loci, due to either insertion or deletion of repeated units (Wheeler, 2005). This change makes them prone to slippage during DNA replication, resulting in a small loop in either the template or the DNA strand. This is usually repaired under normal conditions. However, in the absence of efficient mismatch repair function, the 'loops' may become permanent and alleles of different sizes will propagate at subsequent rounds of replication. If a cancer has silenced or mutated mismatch repair genes, then different sized alleles will accumulate over several generations. MSI is seen when such alleles are formed relatively frequently at microsatellite loci in tumour DNA compared with normal DNA (Wheeler, 2005). DNA mutations in mismatch repair genes are found in all HNPCC (Lynch syndrome) contributing to MSI (Grady and Carethers, 2008, Aaltonen, 1993).

The CIMP Pathway

Epigenetic instability, which is responsible for the CIMP, is also thought to play a part in CRC carcinogenesis. CIMP tumours are characterised by hypermethylation or hypomethylation of oncogene promoters, leading to genetic silencing and loss of protein expression (Marmol et al., 2017, Migliori, 2011). The CIMP pathway is explored further in Section 1.6.3.2.

1.2.3 Symptoms and clinical presentation

Patients with CRC may present in three different ways: 1) suspicious symptoms and signs 2) asymptomatic individuals discovered through routine screening and 3) emergency admission with acute intestinal obstruction, peritonitis or gastrointestinal bleed (Macrae and Bendell, 2017). Most CRCs (70-90% of all cases) are diagnosed after the onset of symptoms, the most frequent of which being: rectal bleeding, abdominal pain, anaemia or a change in bowel habit (Hamilton et al., 2013, Macrae and Bendell, 2017). Change in bowel habits, due to the progressive narrowing of the bowel lumen by tumour, is a more common symptom of left sided cancers. However, the majority of CRC patients who present with iron deficiency anaemia are more commonly found to have right-sided lesions (Kanellos et al., 2004, Thrumurthy et al., 2016). Abdominal pain can occur with tumours arising at all sites, due to partial obstruction, peritoneal dissemination or intestinal perforation. Patients may also present with symptoms of metastatic disease related to the any of metastatic sites of CRC, such as the lung, liver and peritoneum (Macrae and Bendell, 2017)

1.2.4 Diagnosis, staging and grading

Diagnosis

Colonoscopy is still the most accurate test for CRC screening and diagnosing. The National Institute for Health and Care Excellence (NICE) recommends offering a colonoscopy to patients without major comorbidity to confirm the diagnosis of CRC (Poston et al., 2011, NICE, 2014). A biopsy of any suspicious lesions is then taken during the procedure for histological proof of malignancy. Histologically, the majority of cancers arising in the colon and rectum are adenocarcinoma (Macrae and Bendell, 2017). If a cancer is detected, complete visualisation of the whole colon is required pre or post-operatively, as synchronous cancers occur in 5% of patients. Computed Tomography (CT) colonography (CTC) is offered in circumstances where colonoscopy cannot be completed, or where the patient expresses a preference not to undergo invasive diagnostic tests. CTC still requires significant bowel preparation and may not be suitable for patients with extensive comorbidities (Poston et al., 2011, Macrae and Bendell, 2017, Thrumurthy et al., 2016). In cases of major comorbidities or in elderly frail patients, conventional colonic imaging tests may be difficult to

perform due to immobility and poor tolerance of the bowel preparation. In such cases, minimal preparation plain CT is considered (Mahmoud et al., 2017). A plain CT of the abdomen (without prior bowel preparation) has a sensitivity of 88% to 94% for detection of colon cancer (Kealey et al., 2017).

Staging

Staging of the disease and complete visualisation of the colon are required once the diagnosis of CRC is made. All patients are offered enhanced CT scan of the chest, abdomen and pelvis to estimate the stage of the disease. All patients with rectal cancer are offered MRI to assess the risk of local recurrence (Poston et al., 2011) (see Section 1.4.4). The TNM system for tumour (T), node (N) and metastases (M) represents the stages of CRC progression. These stages are then combined into an overall stage definition, which guides therapeutic decisions (Brenner et al., 2014). TNM is the staging system that is most commonly used, as defined by the American Joint Committee on Cancer (AJCC) (Edge et al., 2010). It describes the depth of invasion through the bowel wall, the extent of the involvement of nearby lymph nodes and the presence or absence of distant metastases (Migliore et al., 2011) (Figure 1.3). Clinicians also still use the Dukes' staging system: named after Cuthbert Dukes who, in 1929, proposed a classification designed to represent a step wise progression of local and regional invasion by rectal cancer. The classification has been modified on several occasions to increase its prognostic value, with the commonly used modification produced by Astler and Coller (1954, Mahmoud et al., 2017). Table 1.1 and Table 1.2 describe the AJCC, Dukes' and modified Dukes' staging systems.

Grading

CRC can be graded according to the cancer cell differentiation into four different categories: grade 1 (low grade) if the cancer cells are well differentiated, grade 2 (moderate grade) if the cancer cells are moderately differentiated, grade 3 (high grade) if the cancer cells are poorly differentiated and grade 4 (high grade) for undifferentiated cells (Figure 1.4). Histological grade is considered a stage-independent prognostic factor in CRC, with high tumour grade associated with adverse prognosis (Compton, 2003).

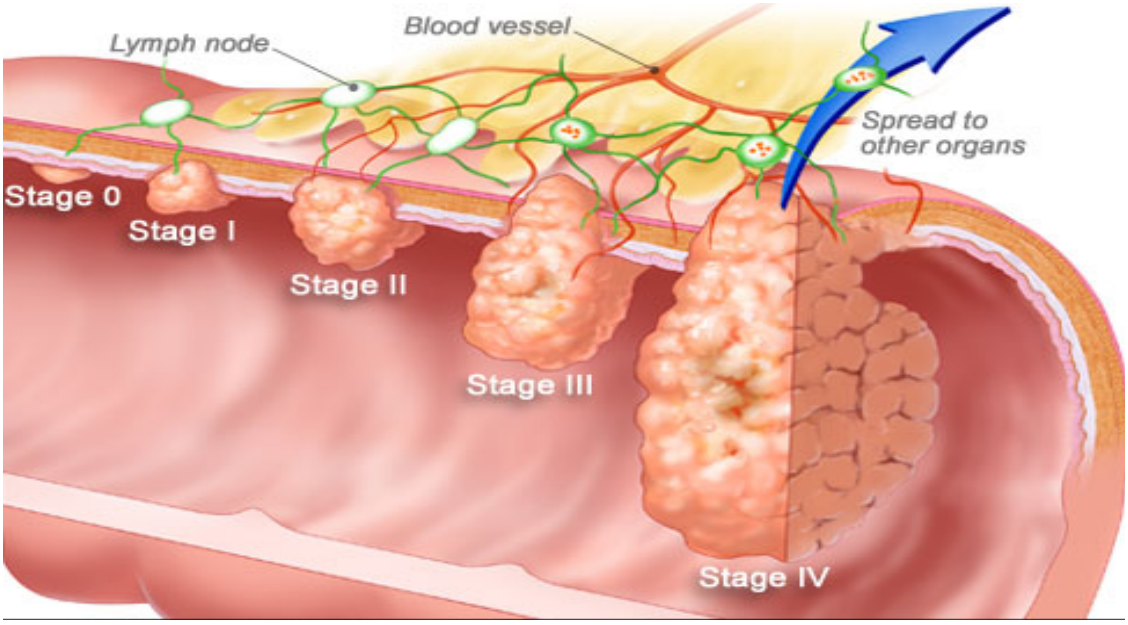


Figure 1.3 The illustrated stages of colorectal cancer. This picture shows different tumour sizes and invasion of anatomical layers of the large intestine. Adapted from the National Cancer Institute, illustrator Terese Winslow (2005).

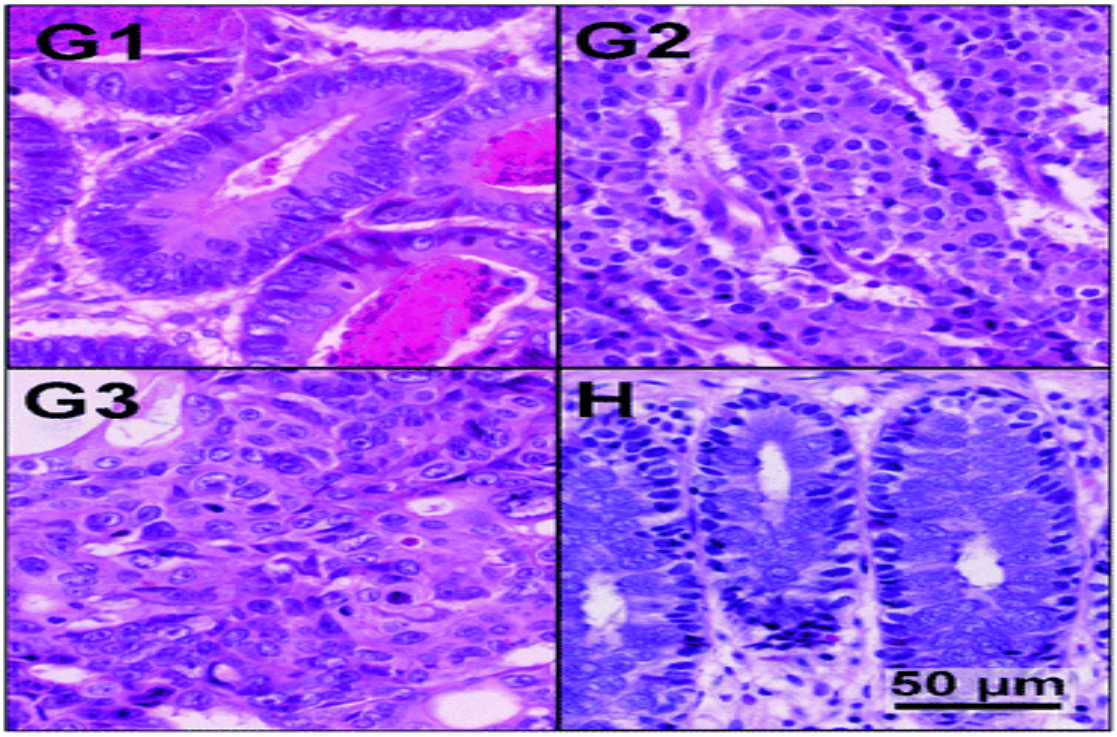


Figure 1.4 Differentiation states of colorectal cancer. Grade 1 (G1), Grade 2 (G2) and Grade 3 (G3) show the grading states during tumour genesis and H represents histological sample of healthy colon with normal cells. Adapted from Kuepper et al. (2016).

T for Tumour: How far the main/primary tumour has grown into the wall of the rectum

Tx: primary tumour cannot be assessed

T0: no evidence of primary tumour

Tis: carcinoma in situ: intraepithelial or invasion of lamina propria

T1: tumour invades submucosa

T2: tumour invades muscularis propria

T3: tumour invades through the muscularis propria into the subserosa or into non-peritonealised peri-rectal tissues

T3a: tumour extends <1 mm beyond muscularis propria 4

T3b: tumour extends 1-5 mm beyond muscularis propria 4

T3c: tumour extends 5-15 mm beyond muscularis propria 4

T3d: tumour extends 15 mm beyond muscularis propria 4

T4: tumour invades directly into other organs or structures and/or perforates visceral peritoneum

T4a: tumour penetrates to the surface of the visceral peritoneum

T4b: tumour directly invades or is adherent to other organs or structures

N describes the extend of the spread to regional (nearby) lymph nodes

Nx: regional nodes not assessed

N0: no regional lymph nodes

N1: metastasis in 1-3 regional (peri-rectal) lymph nodes
N1a: metastasis in one regional lymph node
N1b: metastasis in 2-3 regional lymph nodes
N1c: tumour deposit(s) in the subserosa, mesentery, or non-peritonealised pericolic or perirectal tissues without regional nodal metastasis
N2: metastasis in 4 or more regional lymph nodes
N2a: metastasis in 4-6 regional lymph nodes
N2b: metastasis in 7 or more regional lymph nodes
M describes whether there is spread to other parts of the body – Metastases
Mx: cannot be assessed
M0: no distant metastasis
M1: distant metastasis
M1a: metastasis confined to one organ or site (for example, liver, lung, ovary, non-regional node)
M1b: metastases in more than one organ/site or the peritoneum

Table 1.1 TNM classification of colorectal cancer (Edge et al., 2010).

AJCC Stage	TNM Classification			DUKE'S	MAC*
0	Tis	N0	M0	-	-
I	T1	N0	M0	A	A
	T2	N0	M0	A	B1
IIA	T3	N0	M0	B	B2
IIB	T4a	N0	M0	B	B2
IIC	T4b	N0	M0	B	B3
IIIA	T1-T2	N1/N1c	M0	C	C1
	T1	N2a	M0	C	C1
IIIB	T3-T4a	N1/N1c	M0	C	C2
	T2-T3	N2a	M0	C	C1/C2
	T1-T2	N2b	M0	C	C1
IIIC	T4a	N2a	M0	C	C2
	T3a-T4a	N2b	M0	C	C2
	T4b	N1-N2	M0	C	C3
IVA	Any T	Any N	M1a	-	-
IVB	Any T	Any N	M1b	-	-

Table 1.2 Colon cancer anatomic stage and prognostic groups. According to the AJCC Staging, TNM classification, Dukes' and Astler-Coller classifications**.** (Edge et al., 2010, Astler and Coller, 1954, Mahmoud et al., 2017).

*Modified Astler-Coller classification

1.2.5 Treatment and prognosis

CRC survival is correlated with stage at the time of diagnosis. The 5-year survival rates are 82.9% for early stage disease and 64.4% for late stage disease (Tsikitis et al., 2009). The curative treatment options available at present for high-risk colorectal tumours is complete surgical excision with

adjuvant (postoperative) therapy (Morley-Bunker et al., 2016) or a long course of chemoradiotherapy (CRT). In colon cancer surgery, the tumour and corresponding lymph nodes are removed. The extent and type of surgery is determined by the tumour localisation and the supplying blood vessels. For patients who have undergone a potentially curative resection of colon cancer, the goal of adjuvant chemotherapy is to reduce disease recurrence rate and increase cure rate by eradicating micro-metastases (Rodriguez-Bigas and Grothey, 2017). Despite this, a significant proportion of patients with stage II and stage III CRC suffer recurrence, with risk of recurrence ranging between 15-50% (Brenner et al., 2014). Adjuvant chemotherapy is recommended to patients with stage III colon cancers after surgical resection (Brenner et al., 2014), however, the risk of toxicity may increase in older patients and those with comorbidities, therefore the benefits of the therapy vs the risks should be considered in these cases. A Cochrane review (Figueredo et al., 2008) looking at adjuvant therapy for completely resected stage II colon cancer showed that chemotherapy improves Disease Free Survival (DFS) but had no effect on Overall Survival (OS). Therefore, adjuvant chemotherapy in Stage II is only recommended for patients with high risk of relapse (Brenner et al., 2014) but is usually discussed with individual patients. Most adjuvant chemotherapy involves a combination of several drugs, given intravenously in a specific timeline and order. A six-month course of Oxaliplatin containing regime is usually recommended for the majority of patients (Rodriguez-Bigas and Grothey, 2017). The benefits of Oxaliplatin however are controversial in the elderly and chemotherapy regimens are continuously evolving. Data for the role of neoadjuvant (preoperative) therapy in colon cancer is rare. A recent pilot study showed that preoperative chemotherapy was feasible for patients with locally advanced tumours (FOxTROT, 2012), however further validation in larger cohorts is needed for definitive conclusions. Similarly, adjuvant radiation therapy is not considered a routine component of treatment for patients with completely resected colon cancer.

1.3 THE RECTUM

1.3.1 Anatomy

Anatomical landmarks of the rectum are important for tumour staging, grading and treatment planning. The rectum is the direct continuation of the sigmoid colon and commences at the level of S2-S3. The rectum starts at the recto-sigmoid junction (Mahadevan, 2017) and ends distally at the muscular anorectal ring or junction, which is situated 4 cm anterior to the tip of the coccyx (Jorge and Habr-Gama, 2014). The upper third of the rectum in located intraperitoneally, the middle third is partly located in the peritoneum and the lower third (the ampulla) lies in the extraperitoneal plane (Salerno et al., 2006). The rectum measures 15-20 cm in total length in adults depending upon definitions used. Located in the narrow pelvis, it is surrounded by numerous vital structures such as large vessels, nerves, bladder, internal genital organs and sacrum (Tamas et al., 2017). Fibres from S2-S4 innervate the rectum whilst branches from the inferior mesenteric arteries provide its blood supply (Li and Lai, 2009). The venous drainage of the upper rectum carries blood via the superior rectal vein to the inferior mesenteric vein to enter the portal venous system. The mid and lower rectum drains via the inferior rectal veins, which drain into the internal iliac veins then the inferior vena cava (Mahadevan, 2017). This distinction is important for clinical presentations of metastatic rectal cancer. The rectum is directly surrounded by a layer of fat along its length, known as the perirectal fat. A distinct circumferential layer, known as the fascia propria of the rectum, surrounds the perirectal fat. The fascia propria is an extension of the pelvic fascia and it encloses the rectum, perirectal fat, nerves, blood, lymph nodes and lymphatic vessels (Jorge and Habr-Gama, 2014). The fascia propria with its contents along with the perirectal fat is known as the mesorectum (Mahadevan, 2017). Most of the rectum is surrounded by the mesorectum. This is an important anatomical landmark as the mesorectum or mesorectal fascia forms the circumferential resection margin (CRM) in rectal cancer surgery (MERCURY, 2006) (Figure 1.5).

The visceral fascia that surrounds the mesorectum is reflected on the parietal pelvic fascia at the level of the pelvic floor. Loose connective tissue provides an

interface between the visceral and parietal fasciae, which facilitates dissection and forms bloodless cavities filled with air. This forms the surgical plane of rectal surgery (Salerno et al., 2006) (Figure 1.6).

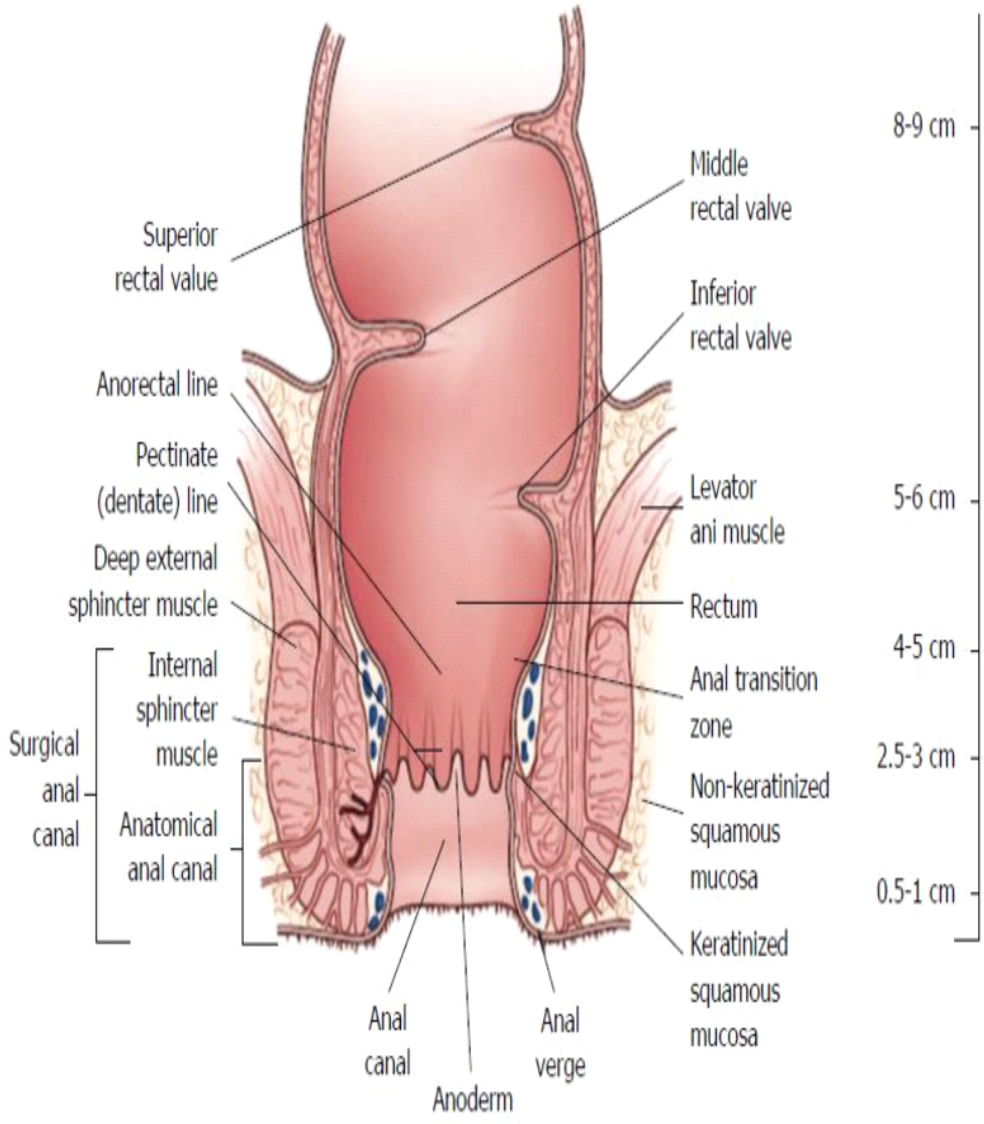


Figure 1.5 Rectal anatomy and important landmarks for rectal cancer treatment. Adapted from Apgar et al (2008).

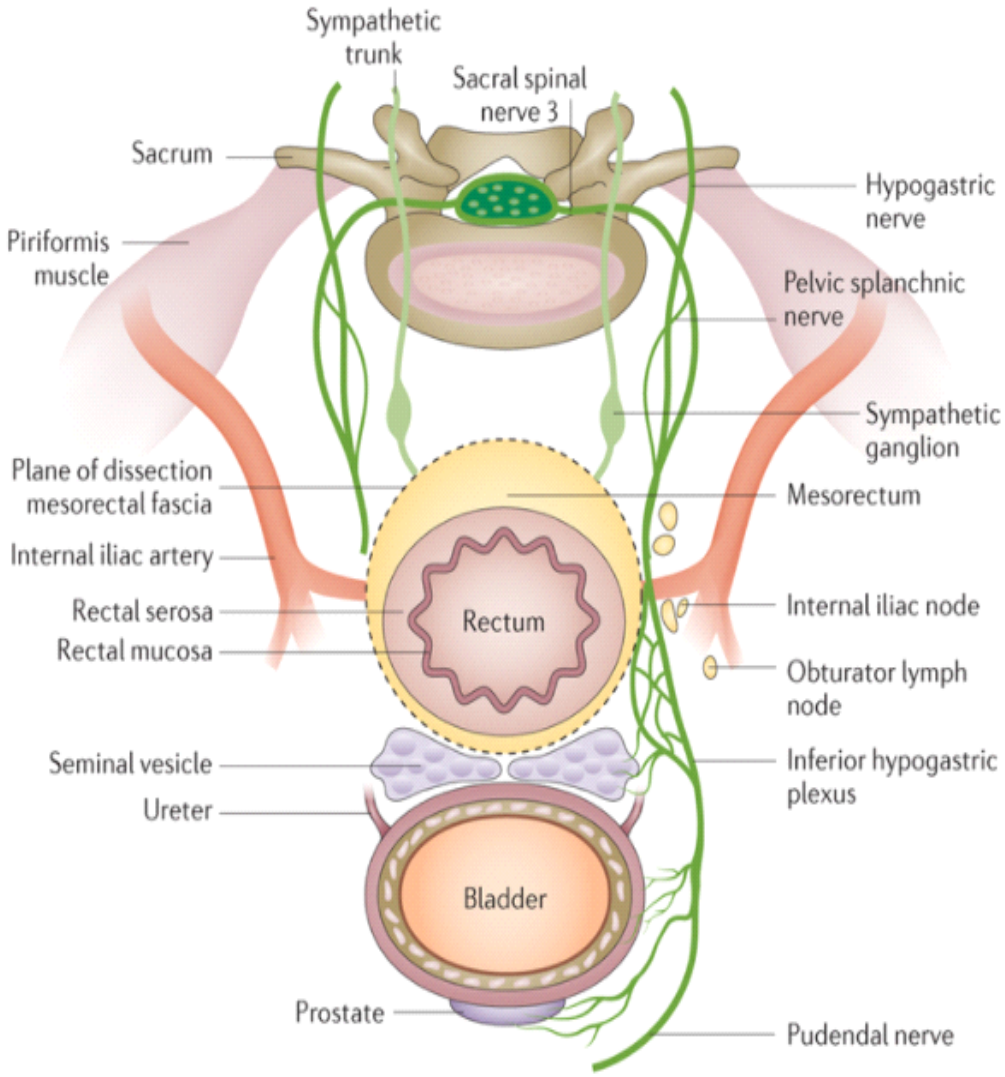


Figure 1.6 Anatomical surgical plane for rectal surgery. Adapted from Kuipers et al. (2015).

1.3.2 Differences between the rectum and colon

The colon and the rectum have different embryonic origins: the rectum arises from the hindgut and the colon from the midgut (Tamas et al., 2017, Li and Lai, 2009). Cancers arising from these two different locations in the large bowel therefore have distinctive features. From an anatomical perspective, the proximal and distal colons are located within the peritoneal cavity whilst the rectum lies within the pelvis, the location of which is not easily accessible (Li and Lai, 2009). Differences between adult proximal colonic mucosa and rectal mucosa have also been reported for the expression of glycoconjugates. Rectal mucosa expresses acidic mucin whilst proximal colonic mucosa expresses neutral mucin (Iacopetta, 2002).

Additionally, the rectum and colon are exposed to different faecal matter and serve different functions. The faecal matter in the rectum is more concentrated than in the colon. Furthermore, as undigested matter travels through the colon it becomes progressively more coated with alkaline mucus (Wei et al., 2004). This means that there are varying levels of pH between the rectum and colon that create varying environments, making the rectum more susceptible to different environmental and risk factors than other parts of the colon.

1.4 Rectal Cancer

1.4.1 Epidemiology

Rectal cancer is defined as a tumour with its lower edge within 15 cm from the anal verge (Brown and Daniels, 2005, MERCURY, 2006). For most epidemiological studies, rectal cancers have been considered as part of CRCs (see Section 1.2). Rectal cancers account for one third of all colorectal malignancies, which in 2008 corresponded to approximately 450,000 new cases worldwide and 15,000 in the UK (Tamas et al., 2017). They are the second most common cancers in the large intestine after proximal colon cancers (Siegel et al., 2017). There is a male predominance for the disease worldwide (30-50% higher than in women) (Tardivo et al., 2005). The risk increases with age, with 70 years being median age at diagnosis in most European countries (Glimelius et al., 2013).

1.4.2 Aetiology

Rectal cancer has largely been studied under the banner of CRC with the two combined as one entity. Hence, its aetiology is similar to that of CRC (Section 1.2.2). The molecular pathways involved in the aetiology of this disease were highlighted in Section 1.2.2.3. The two main genetic syndromes, FAP and HNPCC or Lynch syndrome, constitute the inherited form of CRC and therefore rectal cancer. FAP is associated with mutations in the adenomatous polyposis coli (APC) tumour suppressor genes, whilst HNPCC results from inactivating mutations in DNA mismatch repair genes. In patients with FAP, tumours develop in the distal colon in approximately 60% of cases and in the rectum in 25% (Tamas et al., 2017). As for patients with HNPCC, 55% of tumours are present in the proximal colon and 15% in the rectum.

1.4.3 Symptoms and clinical presentation

Rectal bleeding is the most common presentation of rectal cancer. In later stages, symptomatic patients may present with: tenesmus, rectal pain, diminished calibre of stools, incomplete stool evacuation, cramping, pelvic pain and even obstructive symptoms (Macrae and Bendell, 2017, Fazeli and Keramati, 2015). Adjacent soft tissue organ invasion can occur into the uterus, vagina, anal sphincter complex, bladder, prostate, sigmoid colon or small bowel and rarely into the sacrum (Salerno et al., 2006). Patients suffering with metastatic cancer may present with symptoms related to their sites of metastases. In rare situations, rectal cancer can present as an emergency with intestinal obstruction, acute gastrointestinal bleeding or peritonitis if the tumour perforates the peritoneal cavity. Other rare presentations include fistula formation into adjacent organs, bacteraemia or sepsis (Macrae and Bendell, 2017).

1.4.4 Diagnosis, staging and grading

Diagnosis and Staging

Rectal cancer staging defines the local and distant extent of the disease (Wu, 2007). Exact local staging at the time of diagnosis is essential and is the basis for determining whether a patient undergoes neoadjuvant treatment (Brenner et al., 2014). Similar to CRC, all patients are offered enhanced CT scan of the chest, abdomen and pelvis (see Section 1.2.4).

Endoscopic ultrasonography is accurate for determining the T stage of the disease and is the method of choice for regional tumours or early (T1 or T2) rectal tumours (Wu, 2007). For more advanced rectal cancers, the CRM, the integrity of the mesorectal fascia (MRF) and the remaining lateral and inferior borders of the mesorectum are important landmarks for staging purposes. Although not included in the TNM staging system (Section 1.2.4), local tumour staging in rectal cancer currently includes tumour proximity to the MRF along with depth of tumour penetration and lymph node metastases (Kosinski et al., 2012).

The ability to identify these parameters by dedicated pelvic MRI as well as predicting the degree of extramural vascular invasion (EMVI), was the basis for the Magnetic Resonance Imaging and Rectal Cancer European Equivalence

(MERCURY) trial (MERCURY, 2006). Following the MERCURY study, MRI (thin section, high resolution, phased array coil MRI of the pelvis) became the preferred method for T3 and T4 tumours. Table 1.3 compares the T staging criteria of the AJCC vs MERCURY trial for rectal cancer.

The risk of local recurrence is also assessed using MRI as determined by anticipated resection margin, tumour and lymph node staging which is offered to all patients unless there is contraindication (Poston, 2011), and should be done prior to neoadjuvant therapy as the changes induced by radiation make local staging less reliable (Brenner, 2014).

Grading

The grading of rectal tumours is similar to that of CRCs (see Section 1.2.4). Histopathological examination should include surgical specimen with proximal, distal and circumferential margins, regional lymph nodes and vascular invasion.

	AJCC T CATEGORY CRITERIA	MERCURY TRIAL "T" STAGING CRITERIA
Tis	In situ carcinoma	No corresponding value.
T0	No evidence of viable tumour cells.	No evidence of primary tumour.
T1	Tumour invades submucosa	Tumour invades submucosa. Low signal in submucosal layer or replacement of submucosal layer by abnormal signal not extending into circular muscle layer.
T2	Tumour invades into but not through muscularis propria.	Tumour invades into but not through muscularis propria. Intermediate signal intensity (higher signal than muscle, lower signal than submucosa) in muscularis propria; outer muscle coat replaced by tumour of intermediate signal intensity that does not extend beyond outer muscle into perirectal fat.
T3	Tumour invades through muscularis propria into mesorectal/subserosal fat.	Tumour invades through muscularis propria into mesorectal/subserosal fat. Broad-based bulge or nodular projection (not fine spiculation) of intermediate signal intensity projecting beyond outer muscular coat
T3a	No equivalent category	Tumour extends < 1 mm beyond muscularis propria
T3b	No equivalent category	Tumour extends 1 to 5 mm beyond muscularis propria
T3c	No equivalent category	Tumour extends > 5 - 15 mm beyond muscularis propria
T3d	No equivalent category	Tumour extends > 15 mm beyond muscularis propria
T4		Tumour invades other organs. Extension of abnormal signal into adjacent organ; extension of tumour signal through peritoneal reflection.
T4a	Tumour involves serosal surface.	No equivalent category
T4b	Tumour invades adjacent structures/organs.	See T4

Table 1.3 AJCC vs MERCURY trial for T staging criteria in rectal cancer (Sizer et al., 2006, Beets-Tan et al., 2001, Kosinski et al., 2012).

1.4.5 Treatment and prognosis

1.4.5.1 Introduction

In recent decades, improved surgical techniques, advances in imaging and the introduction of radiotherapy (RT) and chemotherapy have led to improvements in loco-regional control of rectal cancer, but not in survival (Glimelius et al., 2012). The goal in rectal cancer treatment therefore is to optimise DFS and OS whilst minimising the risk of local recurrence and metastases and avoiding treatment toxicity from radiation and systemic therapy (Schrag, 2013, Biondo et al., 2016). Choosing the optimal treatment depends on the risk of local occurrence, the tumour status (T), the nodal status (N) and the surgical CRM (Tamas et al., 2017). The treatment is based on surgery, chemotherapy and RT in different time frames (Petersen et al., 2012, DeSantis et al., 2014, Biondo et al., 2016).

Disagreements and controversies arise when trying to categorise rectal cancer into distinct groups for treatment purposes. Prior to the introduction of MRI staging of rectal cancer and the planning of individualised care through the multidisciplinary team (MDT), the main aim of treatment was to avoid an incomplete resection — an involved CRM (CRM+) — as this is associated with high levels of local recurrence (Burton et al., 2006, MERCURY Study Group, 2006, Quirke et al., 1986). The recognition that preoperative CRT in selected cases and the resultant down-staging of disease through MRI-directed care has led to improvements in OS and a reduction in CRM+ rates. However, selecting the cases that would benefit from CRT, the locally advanced cases, was challenging. In recent decades, there has been an increase in what is considered locally advanced rectal cancer (LARC) with a paralleled increase in improved treatment outcomes. This has created further debate with regards to the nomenclature and treatment of the clinical subgroups of rectal cancer. Currently, treatment decisions are influenced by the four distinct T stage groups (T1, T2, T3, T4) as well as with other relevant factors: tumour height; closeness to the CRM (see Section 1.4.4); clinical N stage; EMVI and nerve invasion. Therefore, it is not possible to define these subgroups based on T and N stages only. Following the MERCURY study (MERCURY Study Group, 2006), the positive experiences of using a dedicated MRI protocol for rectal cancer imaging resulted in the division of rectal cancer patients into three groups for preoperative staging and treatment planning: the good, the bad and the ugly.

1.4.5.2 The good, the bad and the ugly

The good group includes patients with early, clearly resectable tumours, with no bad prognostic factors on MRI, neither for the risk of local or systemic failure (Table 1.4). The slightly more advanced tumours, with MRI features suggesting increased risk for distant metastasis (little challenge to sphincter preservation, limited lymph node metastases), constitute the bad group. Finally, the ugly group included the cases with MRI features suggestive of increased risk for distant metastasis (threatened CRM or CRM+, poor prognostic factors such as lymph node metastases or EMVI) (Kosinski et al., 2012).

Recently, adjusted definitions of the three categories have been widely adopted (Blomqvist and Glimelius, 2008), emphasizing the risk of local failure: a 'good' group has a low risk of failing locally, a 'bad' group has higher risk of failing locally and an 'ugly' group has the worst risk of local failure. Most international groups tend to refer to the 'bad' and the 'ugly' as LARC.

There is an agreement that for the 'good' rectal cancer, surgery is the mainstay treatment with low risk of failure and good survival (Leong et al., 2011). For LARC, curative management focuses on the complete surgical removal of the tumour and prevention of local recurrence and metastases based on surgery, chemotherapy and RT (Petersen et al., 2012, DeSantis et al., 2014). See Figure 1.7 for a summary of the definitions and treatments of the good, the bad, the ugly as described in Blomqvist and Glimelius (2009).

Risk of Local recurrence	**Characteristics of rectal tumours as predicted by MRI**
High	A threatened (<1mm) or breached resection margin OR low tumours encroaching on the inter-sphincteric plane or with levator involvement
Moderate	Any cT3b or greater, in which the surgical margin isn't threatened OR any suspicious lymph node not threatening the surgical margin OR the presence of extramural vascular invasion
Low	cT1 or cT2 or cT3 AND no lymph node involvement

Table 1.4 Risk of local recurrence for rectal tumours as predicted by MRI (NICE, 2014).

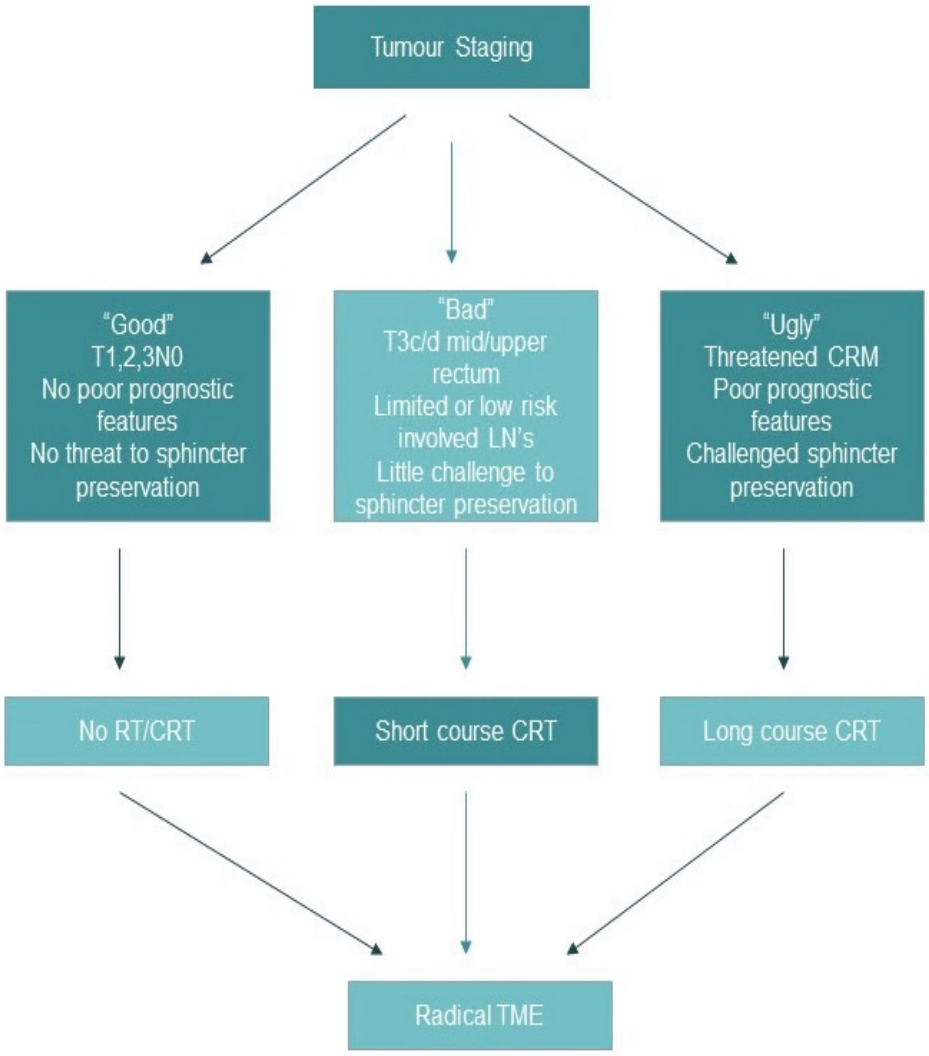


Figure 1.7. The good, the bad, the ugly: definitions and treatment. Adapted from Kosinski et al (2012) and Blomqvist and Glimelius (2009).

1.4.5.3 Surgery

Total Mesorectal Excision (TME) surgery is one of the main components of improvement in current management. For the last twenty years, low anterior resection with TME has been a gold standard surgical treatment for rectal cancer (Dulskas et al., 2016). TME refers to the removal of the rectum together with the mesorectum around the surrounding envelope, the mesorectal fascia. Achieving a clear lateral or circumferential margin is important for the overall

outcome of the procedure. A clear CRM is defined as a distance greater than 1mm between the tumour border and the resection margin, as tumours within 1 mm of the CRM strongly predicts local recurrence and poor survival (Brown and Daniels, 2005). Therefore, an incomplete cancer resection — identified by the involvement of the histopathological CRM (CRM+) — results in increased local recurrence and poor oncological outcomes (Battersby et al., 2015).

Due to the close relationship between the rectum and the pelvic nerves, the bladder, bowel and sexual function are commonly affected following TME (Dulskas et al., 2016). It is therefore important to follow the narrow surgical plane in the pelvis (see Section 1.3.1) to excise the rectum with its lymphovascular supply intact (Brown and Daniels, 2005). The surgical management is therefore technically challenging for the surgeon and local recurrence is a common result of treatment failure (Birbeck et al., 2002).

1.4.5.4 Adjuvant and neoadjuvant therapy

Before the introduction of TME, chemotherapy and RT were adjuncts to surgery, either together or individually. However local recurrence in rectal cancer was common and toxicity of postoperative treatment established. By the late 1970s, the benefits of neoadjuvant chemoradiotherapy (nCRT) became more apparent but so did the risks and side effects. Major randomised trials looking at neoadjuvant treatment demonstrated decreased local recurrence rates with neoadjuvant radiotherapy (nRT) or nCRT vs surgery alone or surgery followed by RT (Kosinski et al., 2012). nRT was also found to have more advantages to postoperative RT, related to tumour response and sparing of normal tissue (Section 1.4.5.5).

The introduction of nCRT for patients with LARC reduced the local recurrence rate from a historical 30% or more, to less than 10% in recent cohorts. It also reduced distant metastases rates and reduced the potential CRM involvement, optimised DFS and OS (Biondo et al., 2016). The main downside of neoadjuvant therapies is the possibility of over-treating patients who may not benefit from it.

1.4.5.5 Radiotherapy and chemotherapy

RT treatment shrinks the tumour in rectal cancer, increasing the likelihood of achieving a pathological clear resection margin and reducing the risk of local recurrence. The target areas in planning for RT treatment are the rectum, anus, sigmoid and mesorectum. RT fields should include the tumour with a 2-5 cm margin, the presacral and internal iliac nodes. External iliac nodes and inguinal nodes are usually included for T4 tumours invading anterior structures and the distal anal canal respectively (Kye and Cho, 2014).

Tumour response depends on sensitivity to RT. Irradiating tissue that is surgically-naïve and therefore better oxygenated may increase sensitivity to RT. Additionally, reducing the volume of tumour pre-operatively through RT may aid the surgical resection and increase the success of sphincter-sparing procedures. Preoperative RT may also avoid postsurgical RT-induced injury to the small bowel (which may be stuck in the pelvis due to post-surgical adhesions). Finally, preoperative RT may additionally target structures that will be resected, increasing the chances of success of surgical anastomoses with a healthy colon (Kye and Cho, 2014, Sauer et al., 2004, Wagman et al., 1998, Kachnic, 2007).

There are two main options for neoadjuvant treatment: a short course of RT or a long course of CRT. Unlike long course radiation combined with chemotherapy, i.e., CRT, short course of RT cannot be safely combined with concurrent doses of systemic chemotherapy. Therefore, the value of choosing a neoadjuvant short course of RT vs CRT is up for debate. In a recent meta-analysis (Chen et al., 2015) looking at neoadjuvant RT for rectal cancer, short course RT was found to reduce the risk of local recurrence compared with surgery alone or surgery with postoperative RT +/- Chemotherapy. The effect was deemed comparable to the local control established with CRT. Short course RT has advantages compared to longer term CRT: it is less expensive, more convenient and has better patient compliance. However, surgery is ideally delayed following short course RT, as cancer cells sterilised by radiation require time to undergo necrosis (Kosinski et al., 2012). On the other hand, a 6-8 week interval following neoadjuvant CRT is standard practice, allowing tumours to undergo necrosis. Nevertheless, a delayed interval to surgery is a reasonable option, especially in elderly patients or patients with comorbidities, who are

often unfit to undergo CRT in locally advanced rectal cancer, prior to surgical resection.

The best regimen for neoadjuvant treatment therefore is not yet fully established, due to lack of consistent differences in terms of local recurrence between different regimes (Biondo et al., 2016). Whilst short course RT (5 Gy per day for 5 days) has been adopted as the standard neoadjuvant treatment in some countries, in others, long course CRT (45-50 Gy over 5 weeks with concomitant chemotherapy) is preferred (Dbeis et al., 2016). Other regimens have used induction chemotherapy prior to CRT, such as in the EXPERT-C trial, by adding Cetuximab in follow up arms (Dewdney et al., 2012).

Follow-up data from important randomised controlled trials looking at different regimes of CRT is yet to be published. Meanwhile, concomitant radiation therapy plus chemotherapy agents that are thought to be radiosensitisers (agents that make the tumour more vulnerable to radiation), such as Fluorouracil (5-FU), Capecitabine or Oxaliplatin is a recognised regime (De Felice et al., 2014, Brenner et al., 2014).

1.4.5.6 Treatment response

The response of rectal cancer patients after nCRT may be associated with oncological outcomes. Preoperative treatment response could serve as an indicator of prognosis and of subsequent response to treatment (Kye and Cho, 2014). The work of Dr Habr-Gamma et al. (2004) in Brazil on the non-operative management of rectal cancer following CRT, led to a desire to identify patients who may respond completely to, thus avoiding surgery, CRT management which is known as 'watch and wait'. Rates of complete clinical response (cCR), disappearance of signs of cancer, and radiological response have been reported as high as 15% in selected U.K. series. Additionally, a group of patients who achieve a complete pathological response (pCR) — no evidence of remaining tumour in the histopathological specimen — has also been identified (Dalton et al., 2012, Renehan et al., 2017).

The role, importance and validity of re-staging rectal cancer after nCRT has been uncertain until recently. The results of the MERCURY study showed that post-CRT MRI assessment of tumour regression grade correlate with DFS and OS, therefore with prognosis. It also gave information regarding the risk of local

recurrence (MERCURY Study Group, 2006, Patel, 2011). The ability to assess local recurrence and prognosis following the initial nCRT enables for a more tailored treatment and ultimately better outcomes (Taylor et al., 2012).
The response of tumours to therapy is often described as tumour regression and measured by tumour regression grades (TRG), which refers to the pathological ratio of residual viable tumour to scar after nRT or nCRT (Kosinski et al., 2012). There is currently no gold standard system to measure TRG, but most examine the tumour specimen histologically. See Table 1.5 for the commonly used pathological systems reported for TRG in LARC. MRI has also been shown to have a role in assessing the response to nCRT (Patel et al., 2011, Battersby et al., 2015) and CT may be used for re-staging.

The high variability in response of patients to the treatment of rectal cancer is attributed to genetic, epigenetic, environmental and patient related factors. The molecular heterogeneity of rectal cancer for example, is believed to be one of the factors responsible for the variability in treatment response among patients with the same stage of cancer (Bettoni et al., 2017). Tumour heterogeneity refers to the differences between tumours of the same type in different patients, or between cancer cells within the same tumour. Both can lead to different responses to therapy, even targeted therapy (Buczacki and Davies, 2014). This may explain why some tumour cells remain present in the patient after completion of cancer treatment. Therefore, single biopsy specimens of primary tumours may not represent the full genetically diverse malignant picture. Similarly, analysis techniques may not be sensitive enough to detect the lower frequency changes in tumour sub clones. Polymorphisms, activation of growth receptor factors and mutations in DNA repair genes may also play a role in the variation in treatment outcomes (Sebio et al., 2015).
Additionally, patient related factors such as sex, age and comorbidities, may influence response to targeted therapies. Certain drugs for example, such as Aspirin and Non-steroidal anti-inflammatory drugs (NSAIDs), are thought to have a potential impact on acute toxicity and pCR. The cyclooxygenase-2 (COX2) pathway is overexpressed in 90% of rectal tumours compared to 20% in colonic tumours (Feng Yin, 2009). COX2 has been hypothesised to inhibit apoptosis, to promote angiogenesis and modulation of cell differentiation and even more to improve cancer aggressiveness and metastasising potential. It is

also associated with chemoradioresistance in rectal cancer (Sakurai et al., 2007, Del Gobbo and Ferrero, 2017). Both Aspirin and NSAIDs inhibit the COX2 pathway.

Other medications may also be responsible in the variation of treatment response: in a multivariate analysis, Katz et al. (2005) found that Statin use is associated with an improved pCR rate after nCRT for rectal cancer. Similarly, Morris et al. (2016) found that the use of ACE Inhibitor and Alpha Receptor Blocker medications among patients with rectal cancer is associated with increased rates of pCR after neoadjuvant treatment.

There are currently no biomarkers to predict treatment response in rectal cancer. However potential candidates are being studied with special focus on methylated genes, which are showing promising results as biomarkers (see Section 1.6.5)

	Mandard	Ryan	*AJCC	Modified Dvorak (pT+pN)
Complete regression	No residual cancer cells (TRG 1)	No viable cancer cells, or single cells, or small groups of cancer cells (TRG 1)	No viable cancer cells (TRG 0)	No tumour cells (TRG 4)
Near complete regression	Rare residual cancer cells (TRG 2)	-	Single or small groups of tumour cells (TRG 1: moderate response)	Fibrosis > 50% of tumour mass (TRG 3)
Moderate regression	Predominant fibrosis with increased number of residual cancer cells (TRG 3)	Residual cancer outgrown by fibrosis (TRG 2)	Residual cancer outgrown by fibrosis (TRG 2: minimal response)	Dominantly fibrotic changes (25-50%) of tumour mass (TRG 2)
Minimal regression	Residual cancer outgrowing fibrosis (TRG 4)	Significant fibrosis outgrown by cancer, or no fibrosis with extensive residual cancer (TRG 3)	Minimal or no tumour cells killed (TRG 3: poor response)	Dominant tumour cell mass (> 50%) with fibrosis or no regression (TRG 1)
No regression	No regressive change (TRG 5)	-	-	-

Table 1.5 Common tumour regression grade systems for locally advanced rectal cancer after treatment. Adapted from Kim et al. (2016).

1.4.6 Differences between rectal and colon cancer

Rectal cancer was usually discussed under the banner of colon cancer in the terms of 'colorectal' cancer. Recently, it has become increasingly clear that the concept of CRC should be divided into proximal colon cancer, distal colon cancer and rectal cancer. Several biological and clinical hallmarks suggest that rectal cancer is different from colon cancer. The rectum and the colon have different embryological origin, anatomy and function (Iacopetta, 2002 and Li and Lai, 2009) (see Section 1.3.2).

Anatomically, colonic tumours extend directly to the serosal surface whilst advanced rectal cancers cannot gain direct access to the peritoneal cavity (Li and Lai, 2009). The rectum is surrounded by numerous vital structures due to its confined location in the pelvis (see Section 1.3.1). Therefore, the local treatment for rectal cancer is more aggressive than for colon cancer. Rectal cancer has a propensity to recur locally. Additionally, it is more likely to metastasise to the lungs whereas tumours in other parts of the colon are more likely to metastasise to the liver (Tamas et al., 2017). This is due to the difference in anatomical venous return between the rectum and the colon (Section 1.3.2).

The staging of colon and rectal cancer are assessed using different methods. The former is assessed with CT scanning but the latter with CT for distant disease staging and endoscopic ultrasound and/or MRI for local staging (see Section 1.2.4 and Section 1.4.4). Historically, the prognosis of rectal cancer was worse than that of colon cancers (Enblad et al, 1988). However, this has changed in the last few years, with worse prognosis more frequently applied to colon cancer patients. This improvement in rectal cancer outcomes is attributed to the evolution of treatment, the standardisation of rectal cancer surgery (TME) leading to a reduction in local occurrence and a steady improvement in survival (Hohenberger et al., 2009, Folkesson et al., 2005).

The incidence of cancer in the colon is similar for both males and females. However, there is male predominance in rectal cancer (30-50% higher than in women) (Tardivo et al., 2005). Other epidemiological differences are also present between rectal cancer and cancers arising in the rest of the colon (see Section 1.2.1 and Section 1.4.1). For example, there has been an increase in the incidence of proximal colon cancer but a decrease in that of rectal cancer in

a number of countries such as Canada, New Zealand, Australia and the United States of America (Tamas et al., 2017). Furthermore, in several epidemiological studies, physical activity was shown to decrease the risk of colon cancer but not of rectal cancer (Samad et al., 2005, Harriss et al., 2009, Wolin et al., 2009). Additionally, behavioural factors, such as physical activity, diet, smoking and BMI, were stronger mediators of risk for colon cancer than for rectal cancer (Doubeni et al., 2012).

There are also genetic differences underlying the two cancers. In contrast to CRC, it has been observed that MSI is rare in rectal cancer whilst the incidence of CIN is high. Yet, compared with colon cancer, the number of mutations detected in rectal cancer is significantly higher (Frattini et al, 2004). Kapiteijn et al (2001) demonstrated that rectal cancers showed significantly more nuclear β-catenin than colon cancers, which is a critical mediator of the WNT signalling pathway - a critical pathway for colon carcinogenesis.

Furthermore, the COX2 pathway, which plays a major role in influencing the development of CRC, is overexpressed in 90% of rectal tumours compared to 20% in colonic tumours (Li and Lai, 2009). Indeed, a recent prospective cohort study of 134,365 men and women concluded that some risk factors, including family history, physical activity and possibly height, differ in their association with colon and rectal cancer (Wei et al., 2004). Therefore, cancers arising from the colon and the rectum have been shown to be different in their biological behaviours, clinical course and genetic mechanisms.

Finally, the molecular heterogeneity of CRC is believed to be one of the factors responsible for the variability of treatment response in patients with the same stage of cancer (Bettoni et al., 2016, Buczacki and Davies, 2014) (see Section 1.4.5.6). This also might account for the difference in behaviour and treatment response between tumours originating in different sites of the colon as well as the rectum.

1.5 Epigenetics

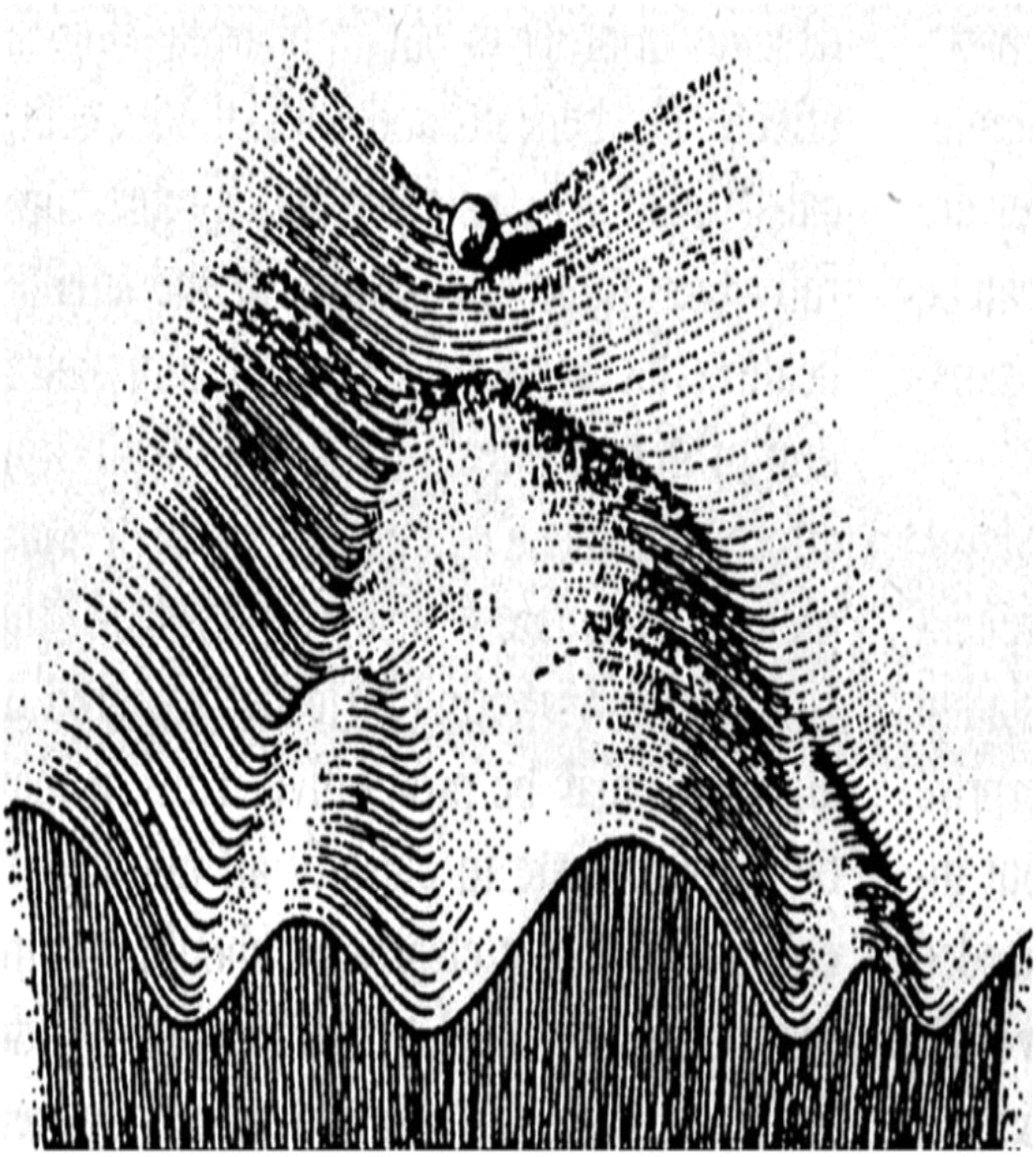


Figure 1.8 Epigenetic landscape. This image shows a cell represented as a marble. The cell is initially phenotypically plural, but becomes increasingly differentiated as it traverses the landscape. This landscape describes the way in which cell-lineage commitments are made during development (Waddington, 1942).

1.5.1 Introduction to epigenetics

1.5.1.1 About the epigenome

One of the challenges of modern science is to understand the mechanisms behind the development of a complex multicellular organism from a single genetic blueprint. A fertilised egg contains a full complement of genetic information, yet, only part of this genetic material is required at all times or by all cells: only half of the ~25,000 protein coding genes within the mammalian genome are thought to be expressed in any cell type (Romanoski et al., 2015). As an organism develops, establishing and maintaining cell-specific genomic regulation becomes necessary: adequate control of genomic functions, such as gene expression and DNA replication, is required for cellular growth and differentiation.

During transcription, a coding portion of DNA (i.e. a protein coding gene) is used as a template (by the enzyme RNA polymerase) to generate messenger RNA (mRNA) molecule. mRNA contains introns, which are sequences that are removed before the mature mRNA leaves the nucleus in a process called splicing. The exons, the remaining regions of the transcripts with the protein-coding regions, are spliced together to produce mature mRNA. After the mature mRNA leaves the nucleus, it is used as a template to synthesise proteins in a process called translation.

The fine-scale control of gene transcription (DNA to RNA) is regulated by the factors known as transcription factors as well as by epigenetic marks. These epigenetic marks also play an essential role in the regulation of developmental processes and their dysregulation has been associated with many diseases, such as cancer (Section 1.5.6).

1.5.1.2 Epigenetics

The term ‘epigenetics’, was first introduced by the embryologist Conrad Waddington in 1942, as the emerging branch of biology that *“...studies the causal interactions between genes and their products which bring the phenotype into being...”* (Waddington, 1942). Waddington famously illustrated his concept as an ‘Epigenetic Landscape’ (Figure 1.8), which shows a cell represented as a marble that is initially phenotypically plural, but becomes

increasingly differentiated as it traverses the landscape (Waddington, 1957). Recent usage of the term epigenetics refers to the study of mitotically heritable, but reversible, changes in gene expression that occur independently of the genomic DNA sequence (Henikoff and Matzke, 1997). Subsequently, epigenetic mechanisms are essential for normal cellular development, tissue differentiation and the long-term regulation of gene function. For a glossary of key epigenetic terms used in this thesis see Table 1.6.

Chromatin (Berger, 2007)	DNA-protein complex that constitutes the chromosomes. The structure of chromatin can be altered through modifications made to the DNA or the proteins associated with it (histones). This facilitates movement between condensed (heterochromatic) and open (euchromatic) states.
CpG island (Gardiner-Garden and Frommer, 1987)	Typically defined as 200 to 500 base pairs (bp) in length with a C and G content of more than 50% and an observed/expected CpG frequency of 0.6.
DNA Methylation (Jaenisch and Bird, 2003)	Addition of a methyl group to carbon 5 of the cytosine pyrimidine ring at palindromic CpG dinucleotides. DNA methylation in certain CpG-rich promoter regions acts to repress gene expression by disrupting the binding of transcription factors and recruitment of proteins associated with chromatin compaction.
Epigenetics (Henikoff and Matzke, 1997)	Heritable, but reversible, regulation of genomic functions occurring independently of the DNA sequence. Primarily mediated by alteration of chromatin structure and DNA modifications.
Epigenetic inheritance (Richards, 2006)	Epigenetic modifications are mitotically heritable and can therefore be maintained across cell division to contribute to cell line establishment. It is less clear whether epigenetic marks are inherited transgenerationally through meiosis.
Histone Modifications (Berger, 2007)	Post-translational, covalent additions made to N-terminal histone tails that modulate chromatin structure. Modifications include acetylation, methylation and phosphorylation.
Genomic imprinting (Davies et al., 2005)	Monoallelic expression of genes in a parent-of-origin specific manner, regulated by allele-specific epigenetic marks established in the germline. Fundamental to normal mammalian development.
Non-coding RNA (Kapranov et al., 2007)	RNA molecules that are not translated into protein that can have structural or regulatory roles.
Nucleosome (Luger et al., 1997)	DNA-histone complex consisting of 147 base pairs of DNA wrapped around eight histone proteins.

Table 1.6 Key epigenetic terms.

1.5.2 Epigenetic mechanisms

1.5.2.1 Introduction

A focus on identifying genetic aberrations has dominated aetiological research in cancer. Over the last decade however, attention has been focused on epigenetic mechanisms: the mechanisms that modify gene expression independently of DNA sequence (Virani et al., 2012, Goel and Boland, 2012). Epigenetic mechanisms are integral to normal cellular differentiation and play a key role in the regulation of gene expression in normal tissue (Avner and Heard, 2001), in genomic imprinting (Morison et al., 2005) and tissue differentiation (Sharma et al., 2010). Once established during development, the epigenetic marks that determine cellular phenotype are mitotically inherited. However, epigenetic alterations can also arise across the life span due to environmental influences (Egger et al., 2004). Therefore, epigenetic modifications can be environmentally induced, tissue specific and can have similar effects to pathogenic mutations: they can silence, increase or decrease the expression of a gene.

In 1982, aberrant epigenetic alterations were discovered in CRC (Feinberg and Vogelstein, 1983). Since, research into epigenetics has revealed a landscape of epigenetic mechanisms controlling gene expression, in both normal and cancerous tissues. Proteins that control gene expression require access to DNA to initiate gene transcription. The accessibility of DNA largely depends on the condensation state of chromatin. An “open” or relaxed chromatin state gives proteins access to DNA for gene transcription, whereas a “closed” or condensed chromatin state prevents gene transcription (Lao and Grady, 2011 and Sharma et al., 2010) (Figure 1.9). The epigenetic landscape is partially a reflection of the factors that determine the condensation state of the chromatin and therefore partly controls gene transcription.

The mechanisms through which epigenetic changes induce their effects include: 1) DNA modification, most commonly DNA methylation; 2) histone modification and 3) microRNAs (Lao and Grady, 2011) (Figure 1.9).

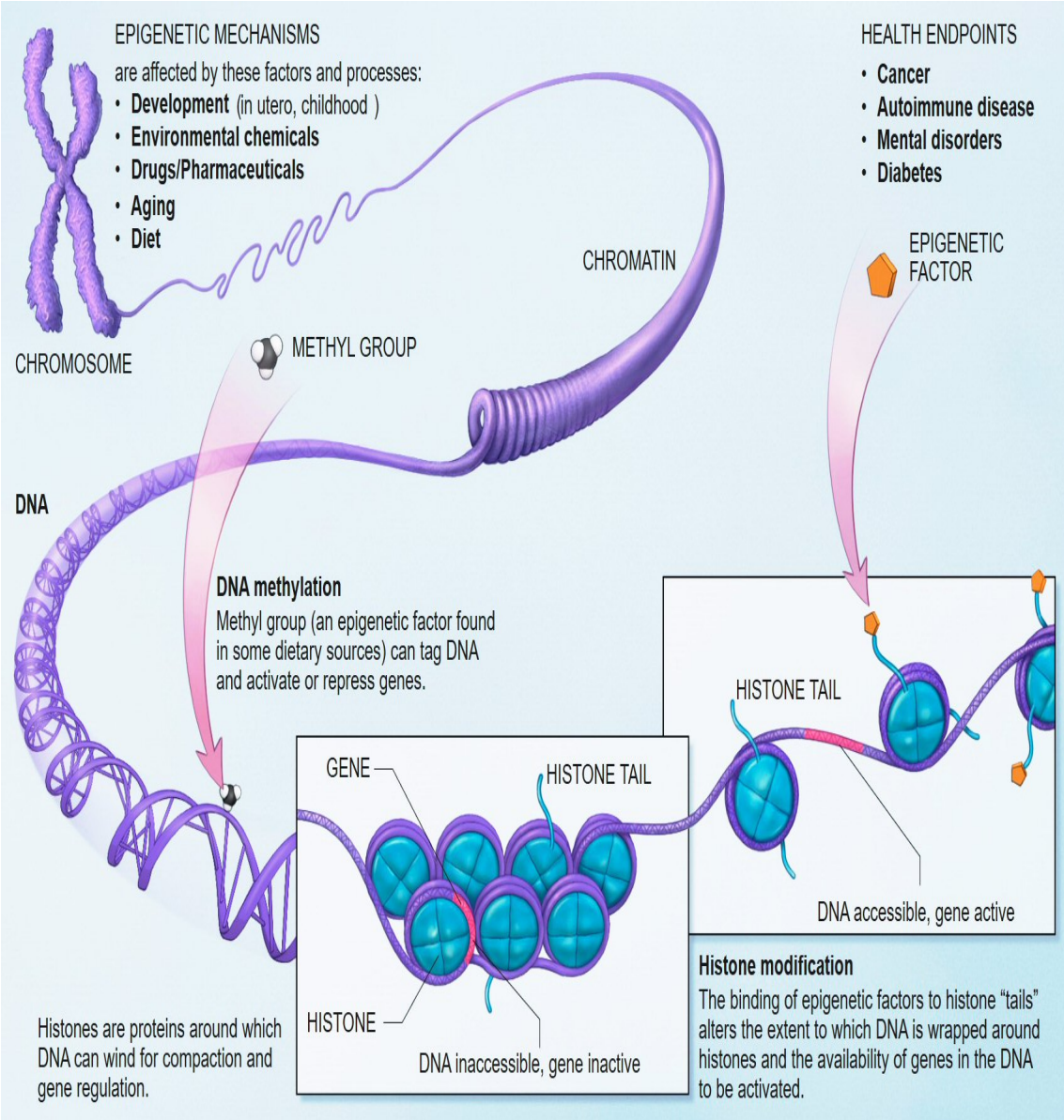


Figure 1.9 Epigenetic mechanisms. The accessibility of DNA largely depends on the condensation state of chromatin. An "open" relaxed chromatin state gives proteins access to DNA for gene transcription, whereas a "closed" or condensed chromatin state prevents gene transcription. Epigenetic mechanisms are affected by several factors including environmental chemicals, drugs and pharmaceuticals, aging and diet. DNA methylation is what occurs when methyl groups tag DNA and activate or repress genes. Histones are proteins around which DNA can wind for compaction and gene regulation. Histone modification occurs when the binding of epigenetic factors to histone "tails" alters the extent to which DNA is wrapped around histones and the availability of genes in the DNA to be activated. Adapted from the National Institutes of Health (2005).

1.5.2.2 DNA methylation

On a molecular level, four bases compose the primary sequence of DNA: adenine, guanine, cytosine and thymidine. However, DNA that has replicated itself in a dividing cell, known as post-replicative DNA, can undergo a covalent modification producing a 'fifth base'. This modification is characterised by the addition of a methyl group to the cytosine ring, forming methyl cytosine, the 'fifth base' in DNA. This reaction uses S-adenosyl-methionine as a methyl donor and is catalysed by enzymes known as DNA methyltransferases (DNMTs) (Figure 1.10).

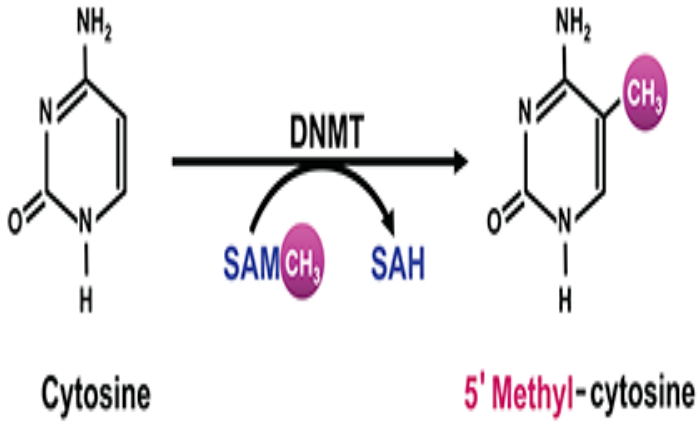


Figure 1.10 Schematic representation of DNA methylation: Methylation converts cytosine to 5'methyl-cytosine via the actions of DNA methyltransferase (DNMT). Adapted from Zakhari (2013).

This stable, non-mutagenic modification is mitotically heritable and contributes to a diverse range of cellular activities, including the regulation of gene expression (Razin and Cedar, 1991, Bird and Wolffe, 1999), the regulation of alternative splicing (Shukla et al., 2011), the control of transcription from alternative promoters within gene bodies (Maunakea et al., 2010), the regulation of genomic imprinting (Sasaki et al., 1992) and the silencing of repetitive DNA elements in the genome (Yoder et al., 1997). This is DNA methylation: the most extensively studied epigenetic mechanism to date. It is the most stable

epigenetic modification modulating the transcription of mammalian genome and will be the focus of this thesis. In vertebrates, DNA methylation predominantly occurs at cytosines that precede a guanine in the DNA sequence (Yi and Kim, 2015). The two base pairs are linked by a phosphodiester bond therefore known as CpG. Mammalian genomes possess conserved regions with a high density of CpG sites known as “CpG islands” (Table 1.6).

CpG islands are often usually sequences greater than 200-500 bases in length and more than 50% of guanine and cytosine in content (Gardiner-Garden, 1987). CpG islands are usually unmethylated if the gene is expressed, and are associated with 60%-70% of gene promoters (Bird, 1986, Williamson et al., 2015) (Figure 1.11). CpG islands have been the focus of much epigenetic research to date. DNA methylation in other genomic regions however, such as inter-genic CpG island shores (Irizarry et al., 2009) and intra-genic CpG islands (Deaton et al., 2011), is increasingly being recognised as functionally important (Figure 1.12). 60% to 70% of the CpG sites in the genome are usually methylated. Methylation is usually associated with the repression of gene expression via physical inhibition of transcription factor and regulator binding (Watt and Molloy, 1988) and through the recruitment of methyl-CpG binding proteins that remodel chromatin into a compact state (Hendrich and Bird, 1998) (Figure 1.9). However, a more nuanced relationship between DNA methylation and transcription, dependent on genomic and cellular context, has recently been revealed (Jones, 2012); DNA methylation at promoter regulatory regions is often associated with repressed gene expression whereas DNA methylation in the gene body is positively correlated with expression (Maunakea et al., 2010).

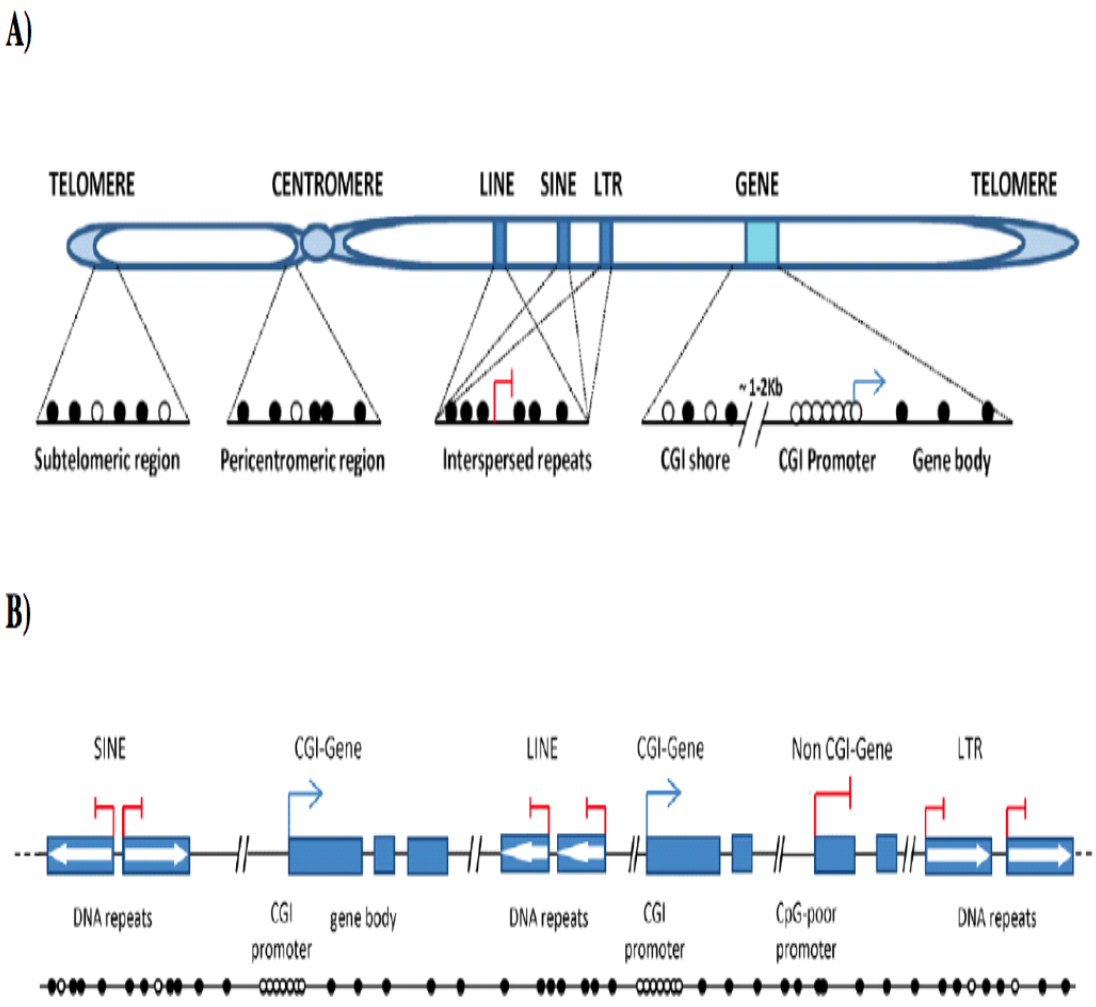


Figure 1.11 Genic location of DNA methylation. A) DNA methylation patterns represented across a chromosome. Repetitive elements and an example of a gene with a CGI-promoter are shown. B) DNA methylation pattern across a genomic region containing three genes and three interspersed repeat classes. Exons are shown in dark blue. The bottom line of each figure represents the DNA methylation state of the DNA; white circles represent unmethylated CpGs and black circles represent methylated CpGs. Blue arrows indicate active transcription and red colour indicates repressed transcription. Adapted from (Zampieri et al., 2015).

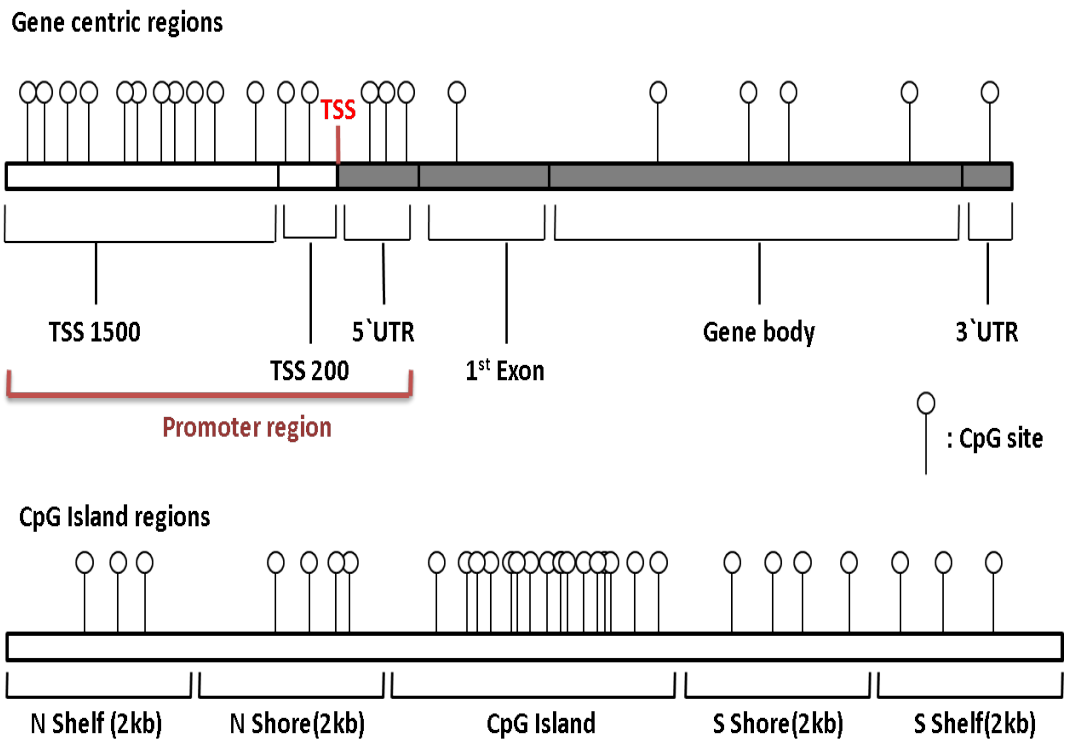


Figure 1.12 Schematic diagram of gene regions and CpG island regions. Adapted from Huang et al. (2015)

1.5.2.3 Histone Modifications

The nucleus of each cell in the genome contains two meters of densely compact DNA. This organisation and compaction of DNA is achieved through the formation of chromatin, the repeating unit of which is a nucleosome (Figure 1.13). Each nucleosome consists of a section of DNA around a positively charged histone octamer that contains two of each of the canonical core histones; H2A, H2B, H3 and H4 (Luger et al., 1997). The N-terminal tails of the histone particles extend out from the nucleosome (Spencer and Davie, 1999). Changes in chromatin structure are mediated in part through post-synthetic, covalent modifications made to these N-terminal. The modifications can influence chromatin compaction state through several mechanisms including; altering nucleosome structure, introducing chemical groups recognised by regulatory or structural proteins, or through the disruption of higher-order chromatin structure (Strahl and Allis, 2000). A full review of histone

modifications is beyond the scope of this thesis, but for a more detailed review see Kouzarides (2007).

These modifications have been hypothesised to form a complex, combinatorial 'histone code' which plays a role in gene expression regulation via alterations in chromatin structure (Berger, 2007). These alterations affect the access of the cell's transcriptional machinery to the DNA in a similar fashion to DNA methylation: in the condensed chromatin state (heterochromatin), in which the DNA and histone proteins are tightly packed, the access to DNA of transcription factors and other co-factors is blocked, repressing transcription. Conversely, an open chromatin state (euchromatin) allows the transcriptional machinery to access DNA and drive transcription (Section 1.5.2 and Section 1.5.3).

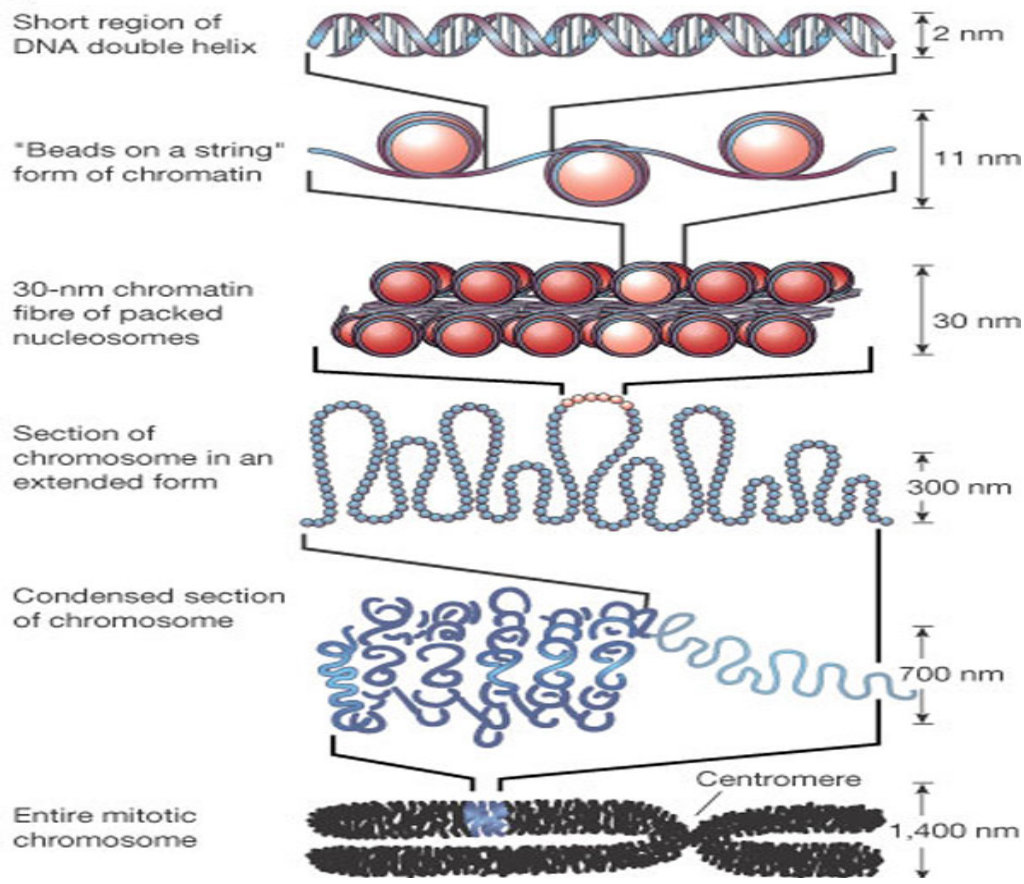


Figure 1.13 The nucleosome, the most basic level of chromatin organisation. The "beads on a string" structure folds into a fibre 30 nm in diameter, which is then folded in to higher-order chromatin structures Adapted from (Felsenfeld and Groudine, 2003).

1.5.2.4 Non-coding RNAs

It is a central dogma in biology that DNA makes RNA and RNA makes protein. However, recent data has shown that most transcripts do not code for proteins. (ENCODE Project Consortium, 2012). Those are known as the non-coding RNAs (ncRNA). Notably, the human genome contains more ncRNAs than protein coding genes (Xie et al., 2014).

Structural ncRNAs are typically required for normal cellular function, and include ribosomal RNAs (rRNAs), transfer RNAs (tRNAs), small nuclear (snRNAs), and telomerase RNA. They can be subdivided according to length: short ncRNA (including microRNA) medium ncRNA and long ncRNA (Prasanth and Spector, 2007). ncRNAs play diverse roles in regulating gene expression, epigenetic processes and defence against viruses. For example, miRNAs contribute to the regulation of gene expression through binding and destabilising target mRNA molecules (Bartel, 2009) and long ncRNAs contribute to the regulation of gene expression through interfering with transcription and inducing chromatin remodelling (Mercer et al., 2009).

1.5.3 Interrogating the methylome

Due to the primary focus on DNA modifications in this thesis, this section will give a brief overview of methods for quantifying DNA methylation.

DNA methylation is erased by standard molecular biology approaches, including the polymerase chain reaction (PCR). Therefore, the detection of DNA methylation requires exposure of DNA to a methylation sensitive pre-treatment prior to DNA sequence analysis, of which the most commonly used is sodium bisulfite conversion (see Chapter 2 Section 2.4.1 for details). Whole genome bisulfite sequencing is considered the gold standard technique to profile DNA methylation, allowing the coverage of ~28 million CpG sites across the human genome. However, the high read depth needed to quantify the small DNA methylation changes makes it economically unfeasible due to a high cost of usage (Ziller et al., 2013).

At present, the Illumina Infinium HumanMethylation450 BeadChip which targets >450,000 (450K) methylation sites, and more recently the Illumina EPIC array are the most commonly used methods for the analysis of genome-wide methylation in humans (see Chapter 2 Section 2.5).

This approach is particularly popular as it represents the best compromise between coverage and cost, and therefore enables DNA methylation to be quantified across the genome in a large number of samples. Further, the widespread use of the 450K array has driven development of many novel bioinformatic methods and pipelines specific for this platform.

1.6 Epigenetic landscape of colon and rectal cancer

1.6.1 Introduction

Epigenetic mechanisms, including DNA methylation, histone modifications and ncRNAs, are known to play a role in carcinogenesis, in the initiation, progression and metastasis of cancer. DNA methylation remains the most studied of these mechanisms for its role in cancer.

DNA methylation was the first epigenetic hallmark to be associated with cancer through altering normal gene regulation and expression (Feinberg and Vogelstein, 1983). It alters gene expression directly by interfering with transcription factor binding and/or indirectly by inducing histone modification leading to a more compacted and repressive chromatin state (Williamson et al., 2015). There are three types of alterations in DNA methylation that have been associated with cancer: hypermethylation, hypomethylation and loss of imprinting. DNA hypermethylation refers to the gain of methylation at specific, normally unmethylated, sites in the genome. DNA hypomethylation refers to the loss of DNA methylation in genome wide regions. Loss of imprinting is defined as the loss of parental allele specific monoallelic expression of genes due to aberrant hypomethylation profiles at one of the two parent alleles (Yi and Kim, 2015). The focus of this review will be on DNA hypomethylation and hypermethylation.

Following the discovery of global DNA hypomethylation in cancer, it was discovered that CpG islands were aberrantly methylated as well. Genome wide studies have shown that 1%-10% of CpG islands are differentially methylated in cancer, amounting to thousands of hypermethylated gene promoters (Costello et al., 2000). Recent studies have discovered a new site for abnormally methylated sites in cancer: the CpG island shores, which refer to areas of less dense CpG dinucleotides located within two kilobases upstream of a CpG island

(Lao and Grady, 2011). Just like CpG islands, the methylation of the CpG island shores is associated with transcriptional inactivation.

1.6.2 Epigenetics of colorectal cancer

The classic view of cancer is that it arises because of the accumulation of mutations in genes or oncogenes, which ultimately causes normal cells to transform into cancer cells. Therefore, genetic aberrations have dominated CRC research and established our understanding of the role of molecular alterations in CRC. Over the last decade, this has led to a shift in research focus on the epigenetic mechanisms in CRC (Virani et al., 2012).

Epigenetic alterations and gene mutations cooperate to drive the progression of normal cells to cancer cells, with the aberrant hypermethylation of genes acting as common molecular mechanism for silencing tumour suppressor genes (Section 1.5). Epigenetic alterations contribute to the heterogeneity of CRC. This is highlighted by the identification of a subtype of CRC with a distinct epigenome and a high frequency of methylated genes (Lao and Grady, 2011) (Section 1.2.2.3). Apart from DNA methylation, histone modifications and non-coding RNAs have also been found to play a role in CRC (Migliore et al., 2011). However, DNA methylation remains the most extensively studied mechanism in CRC. Table 1.7 lists the most frequently methylated genes in CRC, with their known function and the changes in function following aberrant methylation.

1.6.3 DNA methylation in colorectal cancer

1.6.3.1 Background

DNA methylation plays a large role in cancer of the colon. In the average genome of CRC, many more genes are affected by aberrant DNA methylation than by mutation. Furthermore, aberrant DNA methylation is the only mechanism observed for silencing certain genes. When gene expression is altered through DNA methylation, it is usually due to either hypomethylation of global DNA or hypermethylation of promoters.

Nearly three decades ago, Feinberg and Vogesltein (1983) showed extensive

loss of 5'-methylcytosine content in colon cancers compared to normal colon. The global loss of DNA methylation or hypomethylation, appeared to predominantly affect CpG nucleotides found in repetitive sequences of DNA. Later studies showed a causal link between DNA hypomethylation and genetic instability (Costello et al., 2001, Feinberg and Tycko, 2004). Associations between global hypomethylation and aneuploidy in CRC cell lines were also described (Lengauer et al, 1997). Matsuzaki et al (2005) reported that hypomethylation of LINE sequences correlate with loss of heterozygosity on chromosomal loci in CRC. This global hypomethylation was also found in the colon in an age dependent manner as well as in the early steps of the polyp to cancer progression events (Feinberg, 2004, Ehrlich, 2002). Subsequently, global DNA hypomethylation was shown to be an early event in the process of CRC, contributing to cancer formation through inducing genomic instability or loss of imprinting (Lao and Grady, 2011) (Figure 1.14).

In 1999, Toyota and Issa (1999) showed that some CRCs have a high frequency of methylated genes with unique molecular pathogenesis. They named them CIMP cancers. CIMP is characterised by DNA hypermethylation and suppression of key genes, which are differentially methylated between normal and malignant cells and are important in controlling cell growth and transcription (Williamson, 2015). The methylation of the CpG islands within the promoter region is generally correlated with silencing of transcription, however, in CRC it appears that decreased gene expression is only characteristic of certain methylated genes (Hinoue et. al, 2012). Additionally, CpG island shores, which tend to be tissue specific, are altered in CRC (Lao and Grady, 2011).

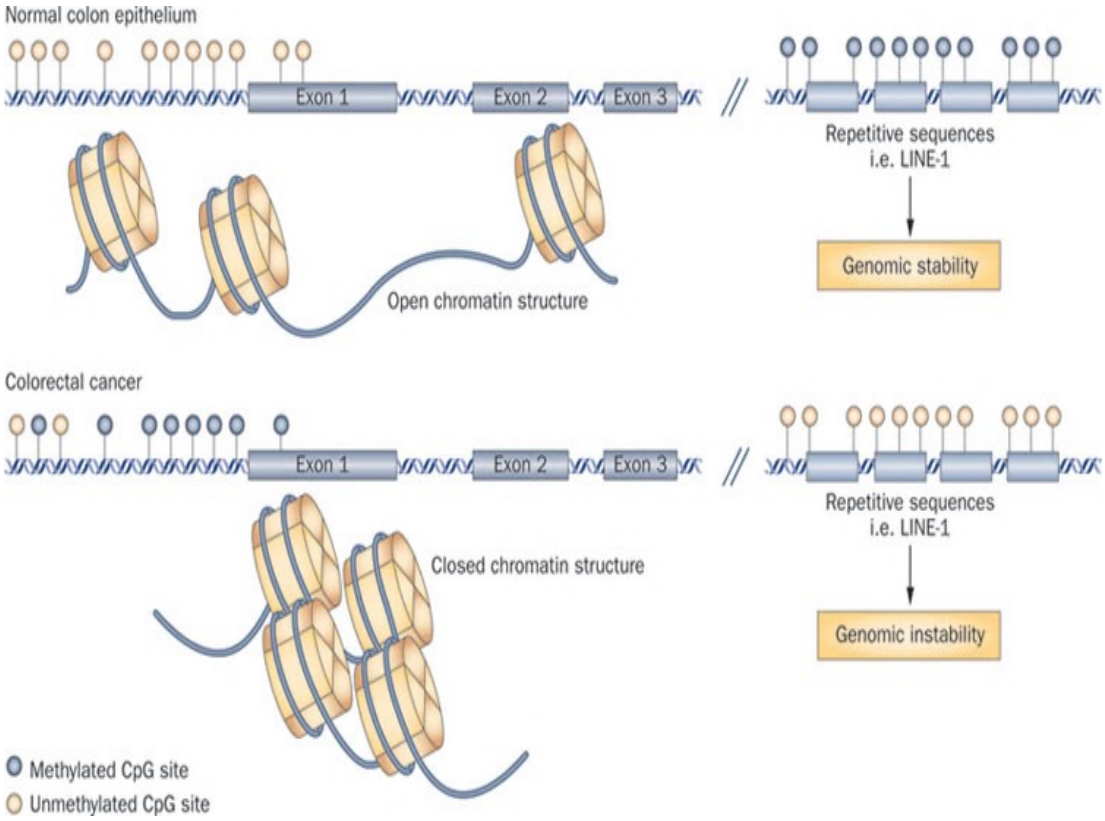


Figure 1.14 Aberrant methylation in colorectal cancer. Aberrant methylation in CRC leads to genomic instability, compared with normal colonic epithelium. Adapted from Lao and Grady (2011).

1.6.3.2 CIN, MSI and CIMP pathways

DNA methylation affects CRC through playing a role in the two main molecular pathways for colorectal carcinogenesis, CIN and MSI, as well through the CIMP pathway, which constitutes a distinct subset of CRC.

CIN is one of the main pathways for colorectal carcinogenesis (Section 1.2.2.2). Global DNA hypomethylation is typically observed in the CIN pathway. A pattern of hypomethylation in pericentromic regions of tumour cells may influence neoplastic progression by making chromosomes more susceptible to breakage. Furthermore, hypomethylation may cause a disruption in the normal structure of genes, resulting in chromosomal instability. These mechanisms may contribute to the chromosomal instability in CRC (Matsubara, 2012).

The MSI pathway (Section 1.2.2.2) is the second main molecular pathway in colorectal carcinogenesis, which occurs due to dysfunction of mismatch genes characterised by the accumulation to numerous mutations targeting DNA

microsatellites (Matsubara, 2012). However, in sporadic CRC with MSI, it is thought that methylation plays an important role in the MSI process. Hypermethylation of the promoter region of hMLH1 for example, causes lack of hMLH1 protein expression and therefore microsatellite instability (Wheeler, 2005). The MLH1 gene is a DNA mismatch repair gene and germline mutations in MLH1 account for around 40% Lynch Syndrome, a hereditary CRC (Boland and Goel, 2010). The loss of MLH1 expression, which occurs where there is biallelic inactivation of the MLH1 genes for example, creates the distinctive mutational signature found in MSI. However, it was also discovered that most of CRCs with MSI that are not attributable to the hereditary Lynch syndrome, known as sporadic CRCs with MSI, have hypermethylation of the MLH1 gene (Weisenberger et al., 2006). This leads to the silencing of MLH1 which leads to the development of MSI. Both types of tumours are characterised by MSI and loss of MLH1, yet germline mutation plays a role in one and DNA hypermethylation in the other making them fundamentally distinct (Boland and Goel, 2010).

Currently, a third subset of CRCs is being recognised and distinguished by the status of DNA methylation at several promoter loci: the CIMP pathway subset. CIMP is characterised by DNA hypermethylation and subsequently a suppression of key groups of genes that are differentially methylated between tumour and normal cells. These genes are normally important for cell growth and survival (Williamson et al, 2015). The panel of marker promoters is used to classify cancers as CIMP positive (CIMP +ve) or CIMP negative (CIMP -ve) (Matsubara, 2012). The first suggested CIMP panel included the following promoter gene regions: MLH1, p16, MINT1, MINT2 and MINT31 (Toyota et al, 1999). Since, several more CIMP panels have been suggested. For example, a commonly used panel as defined by Weisenberg et al includes NEUROG1, SOCS1, RUNX3, IGF2 and CACNA1G (2006). The panel of genes for defining CIMP remains to be standardised and the panel that can be used to classify the CIMP status of colorectal tumours is yet to be determined.

In an extension of the CIMP classification, different CRC epigenotypes have been proposed. These include different classes of CIMP, including CIMP low and CIMP high, which are determined by the proportion of methylated marker loci detected (Lao and Grady, 2011). However, the clinical utility of these designations is hindered, both by lack of a universally accepted definition of

CIMP panel and lack of validation of the accuracy of methylated gene panels which distinguish the different epigenotypes.

1.6.3.3 Initiation, progression and metastasis

Several genes appear to be more commonly methylated in the multi-step process from normal colonic epithelium leading to adenocarcinoma.
At least six genes CDH13, CRBP1, RUNX3, SFRP1, SFRP2, SLC5A8 and two loci MINT1 and MINT31 have been consistently hypermethylated in the transition from normal colonic epithelium to an aberrant crypt focus (Coppedè, 2014, Rashid et al., 2001, Chan et al., 2002, Li et al., 2003, Qi et al., 2006, Kim et al, 2006).

Other examples of genes, such p14, HLTF, ITGA4, CDH1 and ESR1 were found to be frequently methylated in the passage from an aberrant crypt focus to polyp/adenoma (Lao and Grady, 2011). The progression of colon adenomas to adenocarcinomas is also likely to be a consequence of epigenetic alterations, as well as genetic mutations. For example, the hypermethylation of CXCL12, a chemokine ligand, has been found to promote the metastasis of CRC cell lines (Wendt et al., 2006).

Other genes such as TIMP3, ID4 and IRF8, are more frequently methylated in colon cancer and metastases vs adenomas, suggesting that the inactivation of these genes by hypermethylation may lead to a clonal growth advantage (Kim et al, 2006, Kim et al, 2010). For a summary of the commonly cited genes involved in the polyp-> cancer sequence, see Table 1.7 and Figure 1.15.

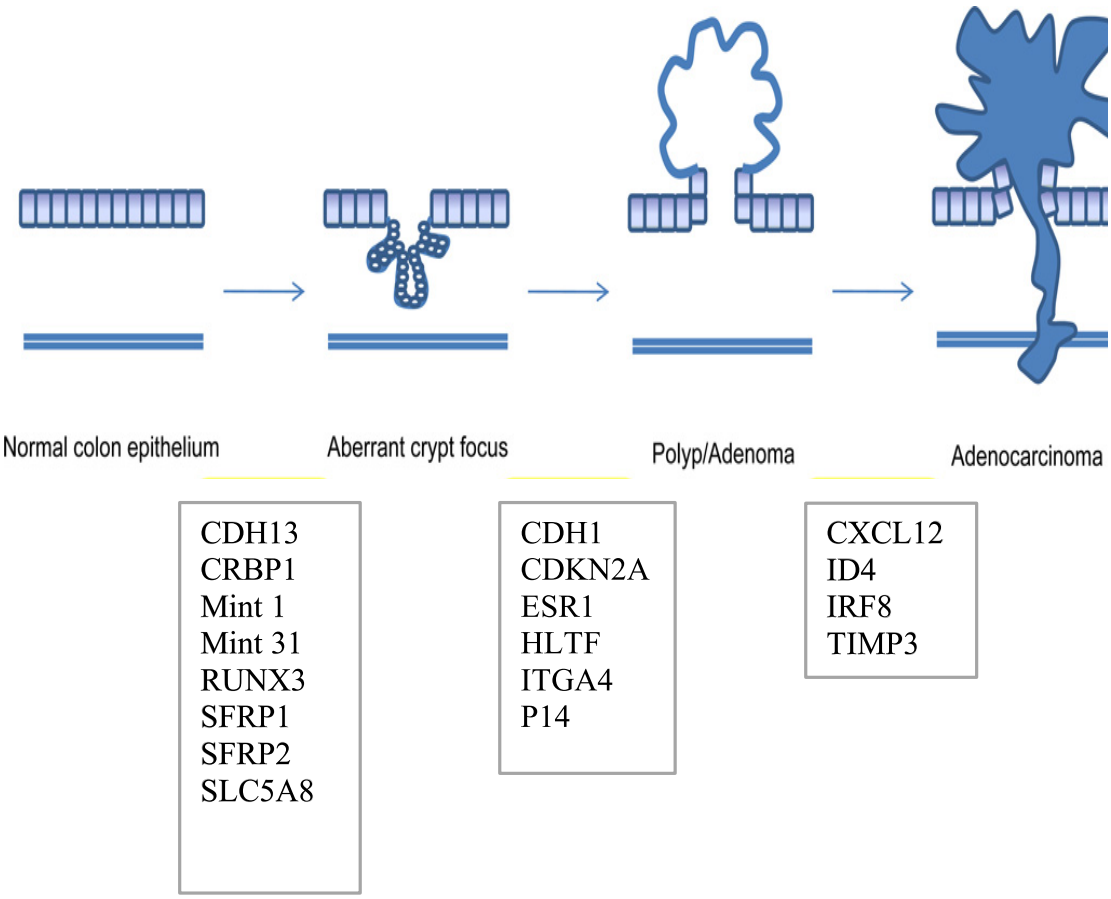


Figure 1.15 Commonly methylated genes involved in the progression of normal colon to adenocarcinoma. Adapted from Lao and Grady (2011)

Gene	Protein	Function	Loss of function effect
APC	Adenomatous Polyposis Coli	Tumour suppressor gene, inhibits Wnt signalling pathway	Increase in Wnt/β-catenin signalling
ALX4	Aristaless-like homeobox 4	Paired-like homeodomain transcription factor	-
CDH1	E-Cadherin	Calcium dependent cell-cell adhesion glycoprotein	Loss of cell adhesion, possible increased Wnt/β-catenin signalling
CDH13	Cadherin 13	Selective cell recognition and adhesion, anti-apoptotic	Increased PI3K/Akt/mTOR signalling, MAPK signalling
CDKN2A (P14, ARF)	Protein 14	Tumour suppressor gene, involved in cell cycle regulation, Inhibits E3 ubiquitin ligase	Decreased p53 stabilisation and activation
CDKN2A/p16	Cyclin-dependent kinase inhibitor 2A	Tumour suppressor gene, regulates cell cycle G1 progression	Increased cell proliferation
CRABP1	Retinol-binding protein 1	Carrier protein for transport of retinol, promotes apoptosis	-
CXCL12	Chemokine (C-X-C motif) ligand 12	Alpha chemokine	Increased tumour cell metastases
DAPK	Death associated protein kinase	Induction of cell death	Interferon gamma signalling, TNF alpha signalling, Fas/APO1 signalling

ESR1	Oestrogen receptor 1	Transcription factor	Loss of oestrogen receptor signalling
HLTF	Helicase-like transcription factor	dsDNA translocase, fork remodelling activity, ubiquitin ligase	Impaired DNA repair
ID4	Inhibitor of DNA binding 4	Transcription factor	-
IGFBP3	Insulin like growth factor binding protein 3	Insulin like growth factor binding protein	-
IRF8	Interferon regulatory factor 8	Transcription factor	Interferon signalling
MGMT	O-6-methylguanine-DNA methyltransferase	DNA repair gene (Repair of alkylation damage)	Increased G→A mutation frequency
MINT1	Methylated in tumour locus 3 (locus not a specific gene)	-	-
MINT31	Methylated in tumour locus 31 (locus not a gene)	-	-
MLH1	MutL homolog 1	DNA Mismatch repair gene	Microsatellite instability
RASSF1A	Ras association domain family 1 (isoform A)	Negative RAS effector, pro-apoptotic, microtubule stabilisation	Increased RAS/RAF/MAP kinase signalling, death receptor-dependent apoptosis
RUNX3	Runt-related transcription factor 3	Transcription factor	Decreased TGF-β/BMP signalling
SEPT9	Septin 9	GTPase, formation of filaments	Impaired Cytokinesis and loss of cell cycle control

SFRP1	Secreted frizzled-related protein 1	Wnt antagonist	Increased Wnt/β-catenin signalling
SFRP2	Secreted frizzled-related protein 2	Wnt antagonist	Increased Wnt/β-catenin signalling
SLC5A8	Sodium solute symporter family 5 member 8	Sodium and short chain fatty acid transporter, suppresses colony formation	-
SOCS1	Suppressor of cytokine signalling 1	Negative regulation of cytokines	-
SPARC	Secreted protein acidic and rich in cysteine (osteonectin)	Inhibition of cell cycle progression, synthesis of extracellular matrix	-
THBS1/TSP1	Thrombospondin 1	Cell-to-cell and cell-to-matrix adhesive glycoprotein	Decreased TGF-β1 signalling
TIMP3	Tissue inhibitor of metalloproteinase 3	Inhibition of MMPs and ADAMs	Increased EGF receptor signalling, TNFalpha signalling
VIM	Vimentin	Stabilising cytoskeleton, member of the intermediated filament family	-

Table 1.7 The most frequently differentially methylated genes in colorectal cancer*. As derived from Migliore et al. (2011), Lao and Grady (2011) and Coppede (2011). *Note, the mechanistic relevance of these changes has not been shown in all instances and it is plausible that many of these differences are a consequence of the cancer rather than have a direct causal role in the oncogenetic process.

1.6.4 DNA methylation in rectal cancer: literature review

Epigenetic aberrations play a signification role in tumour progression and clinical outcomes in rectal cancer (De Maat et al., 2008). However, there are only a few studies that have looked at DNA methylation in rectal cancer specifically, rather than as part of CRC. Most of these studies have analysed individual panels of genes in rectal tumours, previously identified as aberrantly methylated in CRC, and looked at their roles as predictors of clinical outcomes or treatment response.
A literature review on previous studies looking at aberrantly methylated genes in rectal cancer was performed. The list of the genes and CIMP status with relationship to treatment response is highlighted in Table 1.8.

Gene (Reference)	Study Population	Treatment Response	Treatment Given	Other Findings
ATM methylation (Kim et al, 2002)	Human HNPCC-deficient CRC cell lines	Hypermethylation associated with increased response to radiotherapy	10 Gy radiation	Reversal of gene suppression and increased response after AZA treatment
Quantitative DNA 5-methyl-cytosine content analysis (Giotopoulos et al, 2006)	Murine bone marrow cells	Hypomethylation of bone marrow associated with increased radiation sensitivity	3 Gy radiation	-

P16 and hMLH-1 methylation (Hofsetter et al, 2010)	Human CRC cell lines (×4)	Demethylation of markers resulted in enhanced radiation sensitivity	10 Gy radiation	-
MINT loci (De Maat et al, 2010)	251 rectal cancer patients (stages I–III)	Not examined	Primary surgery only	MINT 3 hyper- and MINT 17 hypomethylation predicts reduced risk of recurrence similar to unselected patients undergoing neoadjuvant treatment
Significant hypermethylation in: ESR1, CDH13, CHFR, APC, RARB (Leong K.J. et al. 2011)	51 rectal adenocarcinomas	Not Examined	-	Locus hypermethylation was more prevalent in early- than late-stage disease. Hypermethylation of two or more of a panel of five tumour suppressor genes was associated with localised disease.
TFAP2E methylation (Ebert at al, 2012)	110 human locally advanced rectal cancers	Hypermethylation associated with markedly reduced response to neoadjuvant CRT (10 vs 82 %)	Treatment dose radiotherapy with combination chemotherapy agents	Possible mechanism via WNT signalling pathway
CIMP status +/− (Jo et al, 2012)	150 human locally advanced rectal cancers	No association of pathological	Treatment dose radiotherapy and	Increased risk of distant metastases and poorer 5-year survival with CIMP +

		response to IR with CIMP status	5-FU chemotherapy	
CIMP status H/L/0 (Bae et al, 2013)	168 rectal cancers (stages I–IV)	Not examined	Primary surgery only	CIMP H associated with poor survival
LINE-1 methylation (Benard et al, 2013)	94 early rectal cancers	Not examined	Primary surgery only	Hypomethylation of LINE-1 associated with increased risk of recurrence and poor survival
TIMP3 methylation (Molinari et al, 2013)	74 rectal cancer patients	TIMP3 hypomethylation associated with poor tumour regression	Treatment dose radiotherapy and 5-FU chemotherapy	Several genes including APC found differentially methylated between tumour and normal mucosa
Whole-genome methylation (Gaedcke et al, 2014)	165 locally advanced rectal cancer patients	Not examined	3 cohorts including neoadjuvant 5-FU, Oxaliplatin and radiotherapy	10 differentially methylated regions (DMRs), hypermethylation of which predicts improved disease-free survival
MGMT methylation (Sun et al, 2014)	219 rectal cancer patients	MGMT hypermethylation associated with	Treatment dose radiotherapy and 5-FU chemotherapy	-

		increased tumour regression		
CIMP H/L and CDKN2A methylation (Kohonen-Corish et al, 2014)	381 early rectal cancers	Not examined	Primary surgery for majority of patients	CIMP H or CDKN2A not independently associated with survival. CDKN2A and KRAS mutation associated with poor survival and increased recurrence
CDKN2A, CIMP +/- (Exner et al, 2015)	22 Fresh Frozen Rectal cancer samples	Not examined	Hypermethylation associated with negative prognostic Overall Survival	-
Epigenome wide methylation (Vymetalkova et al, 2016)	25 Rectal cancer patients	Not examined	-	Hypomethylation BPIL3 and HBBP1 hypermethylation TIFPI2, ADHFE1, FLI1 and TLX1 Confirmed by pyrosequencing
CRBP1 hypermethylation (Yokoi et al, 2017)	Six CRC cell lines	Hypermethylation associated with radiosensitivity in rectal cancer	1- 4 Gy radiation	-

Table 1.8 Aberrantly methylated genes in rectal cancer.

1.6.5 Clinical biomarkers

A biomarker is a biological characteristic that can be objectively measured and evaluated as an indicator of normal biological process, pathogenic process, or pharmacological response to a therapeutic intervention (Naylor, 2003). Biomarkers can be used at any stage of a disease and can be associated with its cause (risk biomarkers), onset (diagnostic biomarkers), clinical course (prognostic biomarkers), or response to treatment (predictive biomarkers) (Strimbu and Tavel, 2010).

Epigenetic biomarkers can be used in cancer for diagnosis, prognosis and prediction of response to treatment therefore treatment planning. In fact, epigenetic biomarkers are already in use CRC, and several are being developed for early detection, prognosis and treatment response prediction in CRC as well as breast and lung cancers (Lao and Grady, 2011). Cell free DNA hypermethylation, released from the tumour into the blood and stools, is considered a potential cancer biomarker, due to its potential high sensitivities and specificities in tissue studies in CRC (Schwarzenbach et al., 2011).

Discovery of markers specific for rectal cancers has been hampered by the lack of studies addressing rectal tumours as an independent entity and studies using limited numbers of candidate genes.

1.7 Aim of this study

Key epigenetic features are useful for understanding the molecular subgroups of rectal cancer, the environmental and genetic features that lead to it, as well as predicting the natural history of tumours and the response to different treatment regimens. One of the goals of present and future research is to identify biomarkers that could allow a feasible, cost-effective and non-invasive screening and diagnosis of rectal cancer, as well as to understand which panel of biomarkers can be used to better define patient's prognosis and the best choice of available treatments.

The discovery of DNA methylation patterns and epigenetic markers specific for rectal cancer has been hampered by the lack of studies addressing rectal tumours as an independent entity and studies using limited numbers of candidate genes. Although recent technological advances have facilitated new approaches for identifying DNA methylation changes across the whole genome, to our knowledge, ours is the second study worldwide to use these advances and adopt a genome wide approach in rectal cancer.

Therefore, there are two aims to our study:

Aim 1: To use and validate methods used in CRC studies, for isolating and processing tumour cells from rectal cancer tissues.

Aim 2: To identify methylomic variation associated with rectal cancer across the whole of the genome using the Illumina Infinium 450K array platform.

Chapter 2 - Materials and Methods

2.1 Pre-sample collection

2.1.1 Introduction

The National Institute for Health Research (NIHR) Exeter Clinical Research Facility (ExeterCRF) is a joint venture between the Royal Devon and Exeter (RDE) NHS Foundation Trust and the University of Exeter. It provides staff and facilities to support a variety of mechanistic and translation clinical research projects. One of its key resources is the ethically approved RDE Tissue Bank (RDETB) which was set up to proactively collect and store 'spare' tissue available from routine clinical procedures. Part of the ethical approval of the RDE enables the RDETB steering committee to provide ethical approval on a project by project basis. An application is required to build up a collection and/or access these or previously collected samples. Samples remaining after analysis are returned to the tissue bank for as a resource for future use by others. The RDETB supersedes the original Exeter Tissue Bank (ETB) which collected spare bowel, bladder, kidney and prostate tissue on an ad hoc basis. Access to these stored samples required researchers to obtain full ethical approval for their project prior to application to the ETB. These historic samples were transferred to the RDETB to enable their continued availability for research use. Prior to any sample collection or patient involvement, I completed the Good Clinical Practice course, underwent appropriate training for gaining consent for recruitment of patients into trials. I additionally underwent all appropriate IT, health and safety, administrative and clinical trainings. Template application form and Good Clinical Practice certificate can be found in Appendix (1-2).

2.1.2 Patient and public involvement

Prior to the current study, the clinical members of the research team had regular discussions, during routine clinical care with patients, on potential treatment options and the benefits of each to the individual. From these discussions, it was clear that patients were keen for any research which could help identify the right treatment option for them. This led directly to the development of the current biomarker studies. In keeping with the National Health Service (NHS) Patient Carer and Public Involvement strategy, the ExeterCRF invites user representatives to contribute to the development of various projects within its

portfolio. These individuals have agreed to maintain contact and regular meetings have been established at which researchers discuss the development of current projects within the ExeterCRF. The research team also had access to this group during the project development phase.

2.1.3 Study design

This is a retrospective study using matched samples of rectal cancer and adjacent normal mucosa, from patients with rectal adenocarcinoma, interrogating the differences in genome wide methylation between the two pairs of tissue. Two separate cohorts were included, the historical samples and freshly collected samples as described in Section 2.1.4.

2.1.4 Tissue bank applications

2.1.4.1 Historical samples

Preliminary discussions with the RDETB management team identified a number of existing samples that would be suitable for study. The samples identified were rectal adenocarcinoma, collected between 2004 and 2007, fresh frozen (FF) and stored in the ETB freezers for future use. A total of 58 samples from 29 patients were located. Two matched pairs of fresh tissue samples had been previously collected: a tumour sample and an adjacent normal mucosal sample from each patient. An application to the RDETB for use of these samples was subsequently made and approved (Appendix 3).

2.1.4.2 Freshly collected samples

More recently collected samples were available form a currently approved RDETB project from rectal cancer patients recruited as part of the Risk Stratification for Rectal Cancer Treatment project (RIST-pilot/protocol/Vs1/300914).

2.1.5 Ethical approval and information protection

Ethical approval for the current use of samples was provided by the RDETB

steering committee. All samples held by the RDETB are held in a link-anonymised format. RDETB data and samples are given a unique study/sample ID number. Personal identifiable data is only accessible by RDETB staff with the appropriate data protection training. Project specific consent and paper data collection forms are stored in locked filing cabinets within the controlled access ExeterCRF, for the duration of the projects. Data is transferred onto a project specific RDETB database and stored on a password protected NHS database. Anonymised samples and data were provided to the research team.

2.1.6 Data collection and recording

2.1.6.1 General information

As part of the RDETB data quality procedures, data collected and recorded are screened and reviewed for discrepancies and missing data prior to analysis. All biological samples are stored in line with the Human Tissue Authority (HMGov, 2004) and Medical Research Council guidelines (MRC, 2001) on the handling and storage of human tissues.

Permission is obtained for a member of the RDETB team to access medical notes relevant to the study where appropriate on behalf of the research team. This anonymised data is then provided to the research team as required.

2.1.6.2 Historical samples

Historical colorectal samples collected from 2004-2007 were stored in the ETB database. Clinicians collected all samples during hospital procedures and subsequently fresh froze them in liquid nitrogen. Matched rectal adenocarcinoma samples and adjacent normal mucosa samples were identified, appropriately labelled in tubes and placed in specific locations in the ETB freezers according to date of collection and specific ETB IDs. Consent with the same reference ID numbers as the samples had been obtained from each patient prior to the collection and signed consent forms stored in the ETB database. A trained research clinician, research nurse or ETB officer obtained consent. An application to the RDETB for the use of these historic samples was subsequently made (reference JMRD/Vs1/290116) and approved. The samples were then transferred to the research laboratory freezers for dissection and processing.

2.1.6.3 Freshly collected samples

Participants were identified by clinicians and recruited through the RDETB. Consent with an appropriate RDETB study ID was obtained from each patient prior to the collection of samples. Signed consent forms were stored in the ETB database. Consent was obtained by a delegated member of the RDETB. For each patient recruited, matched pairs of fresh tissue samples were obtained from the tumour and the adjacent normal mucosa. The samples were taken either during rigid sigmoidoscopy in outpatient clinic or prior to surgery at the time of de-functioning stoma formation or rectal resection. Samples were obtained between June 2015 and April 2016. They were then transported immediately in labelled tubes to the NIHR ECRF, which is located on the same premises of the hospital, flash frozen in liquid nitrogen and stored appropriately in the RDETB freezers. The tubes containing the samples were transferred at a later date to the research laboratory freezers for dissection and processing. The consent forms were stored within the RDETB (see Section 2.1.1).

2.1.6.4 Adverse effects:

This was a low risk study. No participants were subject to any adverse events from taking part in the study. A standard procedure is in place for reporting and recording adverse events: The Chief Investigator will be informed of any adverse effects within 24 hours and they will be reported following local NHS RDE SOPs with a copy of any adverse event form stored in the project site file.

2.1.6.5 Study specific sample labelling

Each sample had a unique RDETB ID code corresponding to the patient it was collected from which started with the letters ‘TB’ followed by numbers. Additionally, each sample had a unique sample code, different from the participant ID code. To differentiate between the matched pairs either tumour (T) or normal (N) was added to this sample ID.

2.2 Rectal patients sample processing

2.2.1 Stored historical samples

2.2.1.1 Patient selection

Patients with a rectal adenocarcinoma (high, mid and low rectal tumours) were selected and included in the study. Patients had to be to be over 18 years of age, able and willing to provide consent and not pregnant. A summary of exclusion and inclusion criteria is provided in Table 2.1.

Inclusion Criteria	Exclusion Criteria
➢ 18 years of age	Pregnant
Diagnosis of rectal adenocarcinoma or high-grade dysplasia	Unable or unwilling to provide informed consent
Undergoing treatment with curative intent at the RDE (either surgery or CRT)	
Able and willing to provide informed consent	

Table 2.1 Summary of exclusion and inclusion criteria for patient recruitment.

2.2.1.2 Specimen retrieval

Tissue specimens from 2004-2007 had been previously collected from patients undergoing rectal surgery for rectal adenocarcinoma. Matched samples were collected from the RDE and subsequently stored in the ETB freezers at -80°C.

2.2.1.3 Paraffin embedding and histological assessment

Prior to submitting an ETB application for the retrieval of the tissues, dissection and paraffin treatment of a few initial samples was required to justify the release of all of the required rectal samples. This was to verify the content and quality of the historical samples, and assess the tumour cells for integrity and heterogeneity prior to processing them in the lab. Subsequently, six matched

samples (three tumours and three adjacent normal mucosa) were supplied from the TB freezers, transported in formalin to the RDE Pathology lab where they were paraffin embedded. Thin sections of the fresh frozen paraffin embedded (FFPE) tissues were sliced onto slides. Pathology consultant Dr Ian Chandler reviewed the slides. Dr Chandler confirmed normal colonic mucosa histologically in the normal tissues and the presence of cancer cells in the rectal tumour tissues. Additionally, Dr Chandler estimated a percentage of tumour in each of the three cancer slides reviewed (Figure 2.1). This allowed us to have an approximation of tumour content in each of the cells.

Following the TB application approval, the rest of the historical rectal samples were transferred from the TB freezers to the epigenetics' group -80°C freezers. Sections of these samples were cut, placed in microcassettes and immediately fixed in formalin. The non-dissected part of each sample remained unused and was returned in its original collection tube to the -80°C freezers. The dissection was done speedily to avoid any thawing of the samples, which may affect the quality of the DNA. The microcassettes were then labelled as per TB and study IDs, transported to the RDE Pathology laboratories where they were paraffin embedded together to minimise batch effect. Similar to the test samples, one section of each FFPE sample was transferred onto a slide. The sections were stained with haematoxylin and eosin and reviewed by pathology registrar Dr Sarah Saunders who identified normal colonic mucosa and tumour tissues. Dr Saunders provided an estimate of percentage tumour cells in each cancer sample and confirmed the integrity of normal colonic cells in the normal adjacent mucosal samples. A summary of the histological findings is provided in Section 3.2. All slides and respective paraffin embedded cassettes were labelled with their TB and study code and returned to the epigenetics lab (Figure 2.2).

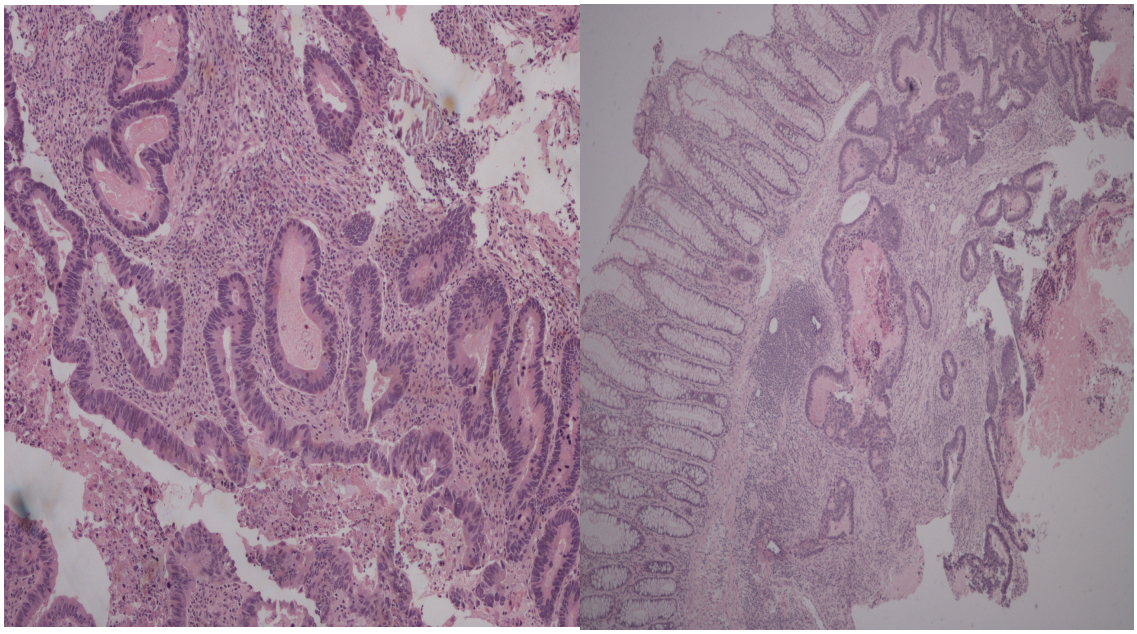


Figure 2.1 FFPE histology slides. Slides from the preliminary histological assessment of the rectal cancer historical samples. They are showing tumour cells, inflamed stroma surrounding the glands and scattered inflammatory cells. Tumour content is different in each of the two slides.

Figure 2.2 Paraffin embedded cassettes of the FFPE tissues. The historical samples lined up for dissection and DNA extraction

2.2.2 Fresh sample collection

2.2.2.1 Patient selection

Clinicians identified patients recruited into the study during clinical visits and treatment, over 18 years of age, with a diagnosis of rectal adenocarcinoma or high-grade dysplasia, not pregnant and able and willing to provide consent. A summary of exclusion and inclusion criteria is provided in Table 2.1.

2.2.2.2. Specimen retrieval

Following appropriate consent, colonic rectal tissues (tumour and adjacent normal mucosa) were collected from surgically removed rectal segments from consecutive patients at the RDE Hospital, Exeter, Devon, during the period of March 2015 to April 2016. Consultant surgeons collected all samples from the operating room, either before or immediately after the surgical resection. In collaboration with the surgical staff, the samples were collected from the operating theatre and transported to the TB facilities in the RILD, which is located on the same premises, either by a TB officer, research nurse or myself. The transfer was done as rapidly as possible to minimise warm ischaemic time and maintain DNA integrity. The samples were immediately snap frozen in liquid nitrogen at the CRF and placed in --80°C freezers. The transfer was done swiftly to reduce DNA degradation.

2.3 Nucleic acid extraction

DNA was isolated from human rectal samples, tumour and adjacent normal mucosal tissue. Two DNA extraction kits were used for the extraction, one for the historical samples which were FFPE and the other for the FF samples recruited for this study. The protocol for each of these extractions is highlighted in the respective sections below (Section 2.3.1 and Section 2.3.2).

2.3.1 Paraffin embedded historical samples

DNA was extracted using the QIAamp DNA FFPE Tissue Kit (QIAGEN GmbH, QIAGEN Strasse 1, 40724 Hilden, GERMANY) optimised for purification of DNA

from paraffin tissue (Figure 2.3). The procedure consisted of six steps summarised in Table 2.2 and briefly described below:

Deparaffinisation: The paraffin tissue was removed from the cassette using a sterile scalpel, placed on a dish where excess paraffin was dissected off the sample block. The sample was then cut into 5-10 very small sections and immediately placed in 1 ml of xylene. This was vortexed for 10 seconds then centrifuged for 10 minutes then washed with ethanol to remove excess xylene. DNA isolation was undertaken as described in the manufacturer's protocol with a minor adjustment - the xylene wash was repeated again, as outlined above, before proceeding to the next step to remove the excess paraffin that was present in the tubes following the first wash.

Lysis of the samples: 180µl of lysis buffer ATL and 20µl of proteinase K were added to the samples, mixed thoroughly and homogenised using a stator-rotor homogeniser. The tubes were incubated at 56°C overnight rather than the one hour standard incubation period, as colonic tissue is hard to lyse. During the 24-hour period, the samples were intermittently vortexed and defragmented using a pestle and mortar, to encourage complete lysis.

Heating: The samples were then incubated in 90°C to reverse formaldehyde modification of nucleic acids and inactivate proteinase K.

Binding of DNA: The DNA was washed in 200µl of Buffer AL and 200µl of Ethanol, and the entire lysate was transferred to the QIAamp MinElute column, which was placed in a collection tube.

Washing the DNA: Residual contaminants were washed away using 500µl of AW1 and 500µl of AW2, the wash buffers, in two distinct steps.

Eluting the DNA: Pure concentrated DNA was eluted from the membrane by adding 40µl of the buffer ATE in total, in two washing steps. The final product consisted of 40µl of DNA.

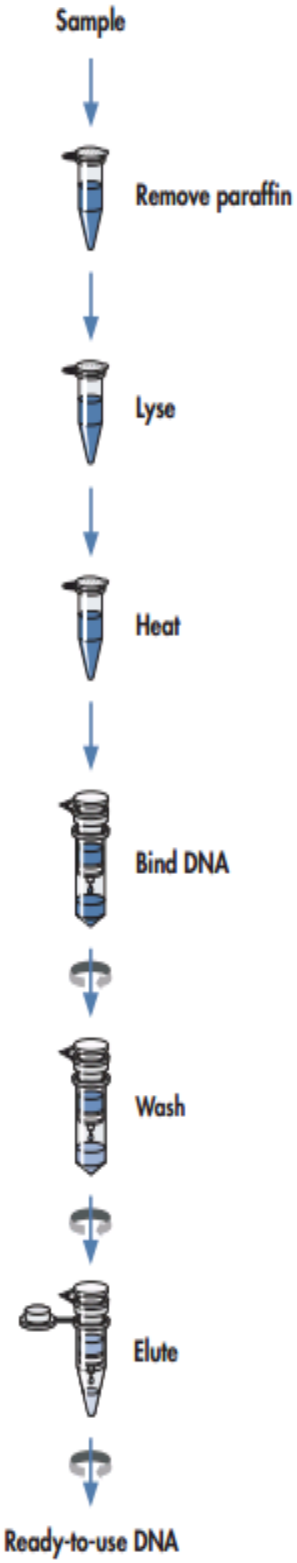


Figure 2.3 DNA extraction from FFPE tissue: Procedure overview
Adapted from the QIAamp DNA FFPE tissue handbook 2012 (QIAamp, 2012).

Remove paraffin	Paraffin is dissolved in xylene and removed
Lyse	The sample is lysed with proteinase K under denaturing conditions
Heat	Incubating the samples at 90°C to reverse formalin crosslinking
Bind	DNA binds to the membrane in the column and the contaminants flow through
Wash	Residual contaminants are washed away
Elute	DNA which is pure concentrated is eluted from the membrane

Table 2.2 The QIAamp 6 step procedure for DNA extraction from FFPE tissues. Adapted from the QIAamp DNA FFPE tissue handbook 2012 (QIAamp, 2012).

2.3.2 Fresh frozen samples:

DNA was extracted using QIAGEN AllPrep DNA/RNA Mini Kit (50) (QIAGEN GmbH, QIAGEN STRASSE 1, 4072 HILDEN, GERMANY). The procedure consisted of four steps summarised in Table 2.3 and Figure 2.4 and briefly described below:

Lysing and homogenising: An appropriate size section (approx. 3mm cube) was cut from each fresh frozen sample on dry ice. Lysis buffer RLT was added to each tube and the tissue initially disrupted with a pestle and mortar then lysed and homogenised with a rotor stator homogeniser. The time spent homogenising the samples was longer than suggested in the kit due to the dense fibrous nature of colonic tissue.

***Binding DNA*:** The homogenised lysate was transferred to the AllPrep DNA Spin column placed in a collection tube and centrifuged for 30 seconds at 10,000 g. This allows the binding of the DNA onto the membrane of the column.

The spin column is then placed in a new collection tube.

Washing DNA: 500µl of the wash buffer AW1 is added to the spin column. The column and collection tubes are then centrifuged at full speed for 2 minutes to wash the spin column membrane. The same step is repeated with AW2, a second wash buffer.

Eluting DNA: The Allprep DNA spin column is placed in a new collection tube and 25µl of buffer EB (preheated at 70°C) is added directly to the spin column membrane. Following a few minutes of incubation, the spin column and collection tube are centrifuged for 1 minute at 10,000g. This procedure is repeated with another 25µl l of buffer EB. The final product consists of 50µl of pure eluted concentrated DNA in the collection tubes.

Lyse	The sample is disrupted then lysed
Bind	DNA binds to the membrane and the contaminants flow through
Wash	Residual contaminants are washed away
Elute	DNA which is pure concentrated is eluted from the membrane

Table 2.3 The QIAGEN Allprep DNA/RNA mini kit (50) 4 step procedure for DNA extraction from FF tissues. Adapted from the QIAGEN Allprep DNA/RNA mini kit handbook 2012 (QIAGEN, 2012)

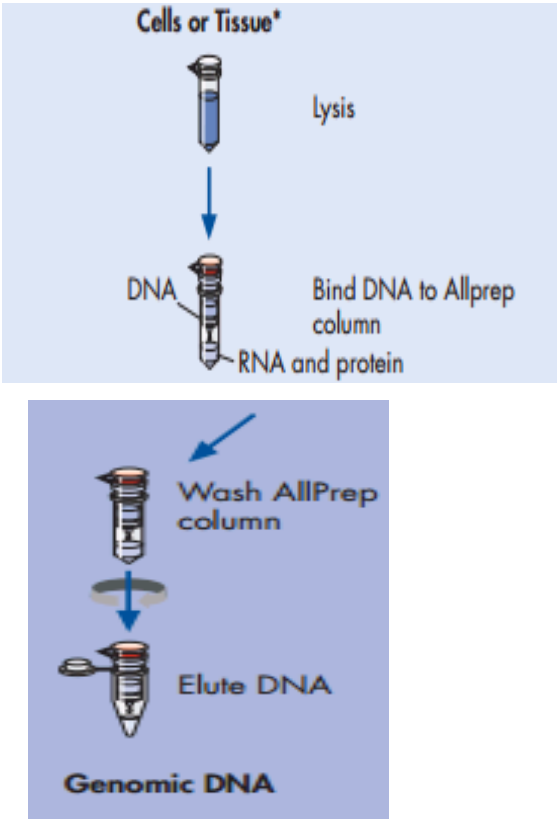


Figure 2.4 DNA extraction from FF tissues. Procedure overview (QIAGEN, 2012)

2.3.3 Agarose gel electrophoresis

Agarose gel electrophoresis is a commonly used technique to separate DNA fragments by their size. DNA is negatively charged at neutral pH due to its phosphate backbone. Therefore, when an electrical potential is present, the DNA molecules will move towards the positive pole in the agarose gel. This gel consists of Tris-borate EDTA (TBE) buffer solution (maintains the adequate pH and salt concentration) with 0.75% to 2% agarose added. The agarose forms a porous lattice in the solution, which allows the DNA to slip through the holes as it travels towards the positive poles. The rate at which DNA travels in the gel is determined by the fragments' size: shorter molecules move faster and migrate further. The separated DNA is then viewed using a stain, most commonly and in our case Ethidium Bromide, which integrated in to the structure of the pre-stained DNA allowing it to be visualised under UV light. UV light is absorbed by the Ethidium Bromide and re-emitted as visible light. This allows the fragments of the DNA to be observed. The DNA ladder or molecular weight marker is a set

of DNA fragments of a known molecular size that are used as a standard to determine the sizes of the unknown fragments placed in the gel. This is usually added first to the gel allowing for approximate assessment of the size of DNA fragments. The application of agarose gel electrophoresis includes assessing the quality and quantity of DNA and determining their size. This technique was used in this thesis to assess the quality of genomic DNA extracted and inspect PCR amplification products (Section 2.4.3). Extracted DNA was run on a concentration of 1.5% and PCR products was run on higher concentration using a 2% gel due to the expectation of larger DNA fragments. The gel was then assessed for the presence of DNA bands, the absence of 'smear' of degraded DNA indicating high quality DNA appropriate for downstream application (Figure 2.5). The protocol used for the production and running of an agarose gel is outlined in Table 2.4.

1. 1.5 g of agarose was added to 100 ml of 1 x TBE buffer
2. The solution was then warmed in a microwave, with intermitted mixing, until the agarose dissolved.
3. 1 µl Ethidium Bromide was then added to the mixture once the solution cooled at room temperature and the whole lot mixed thoroughly.
4. The gel was poured into a tank and combs added. This was left to set for about 25 minutes at room temperature.
5. Once the solution set, the gel was placed in a solution of TBE buffer in a gel tank.
6. The combs were removed gently, leaving wells in the gel.
7. DNA ladder was added to the first well to allow estimation of DNA fragment sizes.
8. 5µl of DNA from samples were added subsequently to other wells with 2µl of orange G loading buffer.
9. The gel was run at 110 volts for 1 hour approximately.
10. The gel was visualised under UV light and photographs were taken.

Table 2.4. Agarose gel protocol

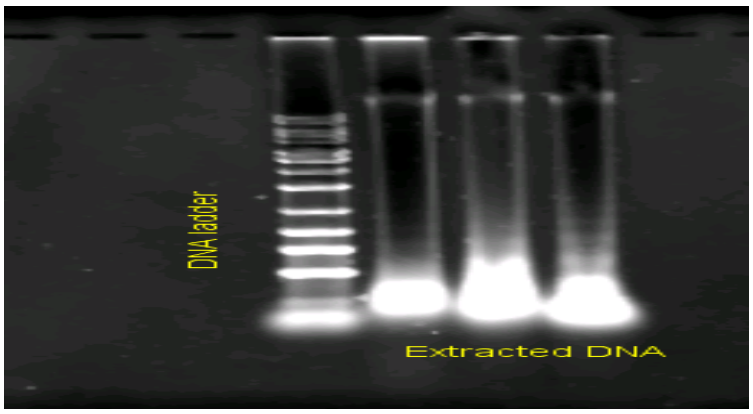


Figure 2.5 Agarose gel (1.5%) used to assess quality of extracted DNA from FF and FFPE samples. An example of 3 samples that show clear bands, absence of smear of degraded DNA indicating successful DNA extraction suitable for downstream applications.

2.3.4 Spectrophotometry

A spectrophotometer was used to assess the quality and quantity of DNA extracted. It is a device that consists of two instruments, a spectrometer for producing light and a photometer for measuring the intensity of the light wavelength. The DNA liquid is placed between the spectrometer beam and photometer. The photometer delivers a voltage signal, which changes as the amount of light absorbed by the liquid changes. DNA can absorb light at 260nm. RNA also has a great absorbance of light at 260nm and amino acids present in protein absorb light at 280nm, both considered contaminants in DNA. Furthermore, strong absorbance around 230nm can indicate that salts or organic compounds are present in the purified DNA. Therefore, a ratio of 260nm to 230nm can help evaluate the level of salt carried over in the purified DNA and a ratio of 260nm to 280nm the level of protein.

A NandoDrop (ND-1000 Thermo Scientific, USA) was used to determine the concentration (ng/µl) and purity of extracted samples. Subsequently, DNA was considered pure if the 260/280 ratios were between 1.7-2.0 and the 260/230 ratio was between 1.8 – 2.2 (Figure 2.6). Lower measurements would indicate the presence of co-purified contaminants such as protein.

Following quantification, an aliquot of each sample was diluted to a concentration of 25 ng/µl and stored at -20°C. Stock sample solutions were preserved at -80°C.

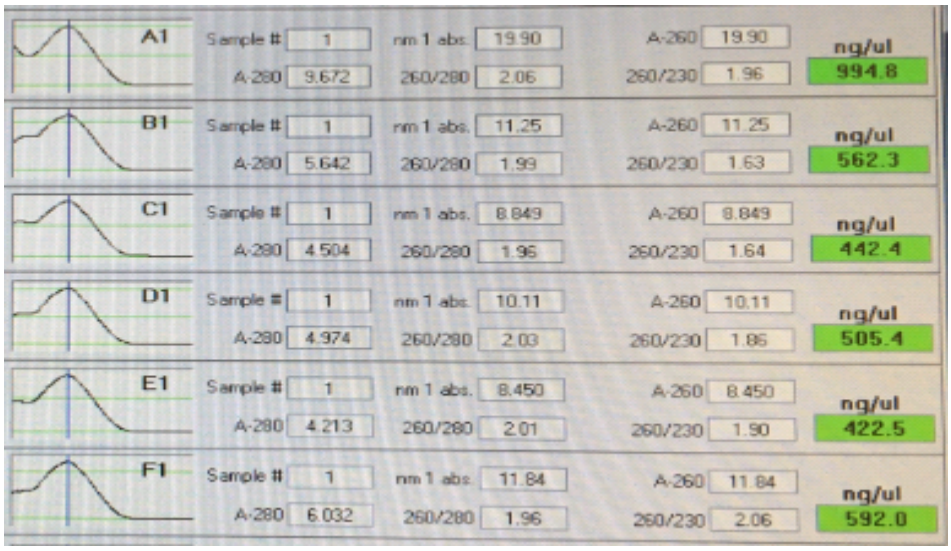


Figure 2.6 Nanodrop samples. Six nucleic acid samples with a characteristic profile, ratios, and concentration.

2.3.5 Dilutions

Manual Dilutions

All samples with a very high concentration of DNA (>500 ng/µl) were diluted in TE buffer to approximately 100 ng/µl in preparation for robotic dilutions

Robotic dilutions

Using an Eppendorf Epmotion Robot all samples were diluted to an exact level of 50 ng/µl prior to Sodium Bisulfite conversion, the next step.

2.4 DNA modification for methylation analysis

2.4.1 Bisulfite modification of DNA

Commonly used sequencing technologies are unable to distinguish between methylated and unmethylated cytosine residues due to their similar base-pairing characteristics. Sodium Bisulfite ($NaHSO_3$) is a chemical that alters the sequence of DNA. It converts unmethylated cytosines to uracils through deamination of these cytosines. Uracil is later replaced by thymine during

downstream procedures. Methylated cytosines are protected from deamination and therefore remain as cytosines (Figure 2.7) (Wang et al., 1980, Paul and Clark, 1996). This is necessary as sequencing technologies are unable to distinguish between methylated and unmethylated cytosines due to their similar base pairing characteristics. A range of technologies can be subsequently used to quantify this genetic alteration allowing determination of DNA methylation status through comparison of expected and observed DNA sequences (Frommer et al., 1992).

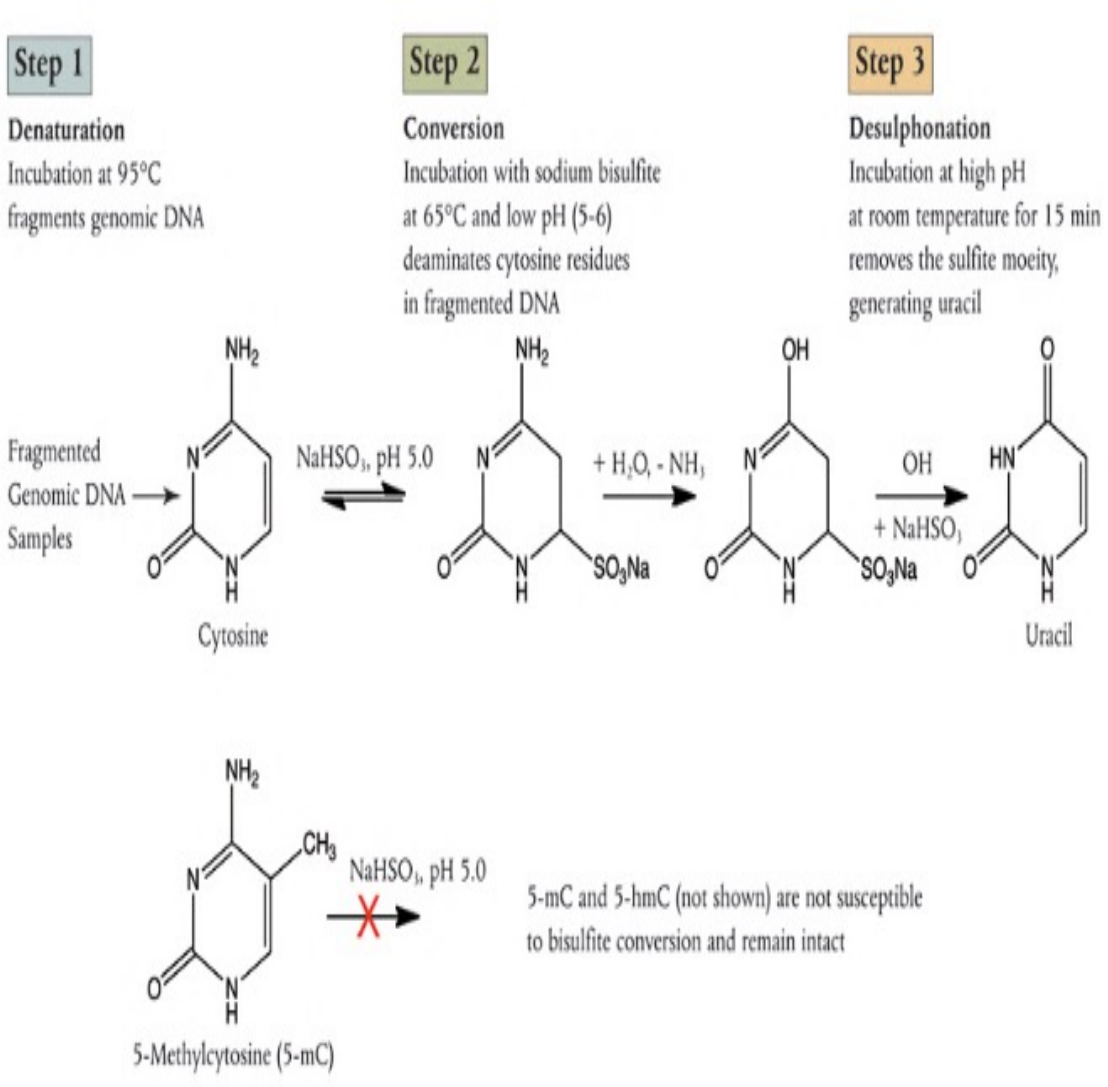


Figure 2.7 DNA sodium bisulfite treatment. Taken from New England Biolabs (Ipsqhich, MA, USA) webpage (New England Biolabs, 2016) (www.neb.com).

Sodium Bisulfite conversion was carried out on the genomic DNA samples, FF and FFPE, at the same time. 500ng of DNA from each sample was treated with sodium bisulfite using the EZ-96 DNA Methylation-Gold Kit (Zymo Research,

Irvine, CA, USA), according to manufacturer's instructions. In brief, the *'CT'* conversion reagent was added to 20µl of each DNA sample in a conversion plate. This was then heated under different temperatures. 400µl of Binding buffer was then mixed followed by a wash buffer. 200µl of desulfonation buffer was added, followed by a wash buffer. Finally, the 30µl of treated DNA was eluted for each sample. Figure 2.8 shows an overview of the experimental procedure. The detailed steps of the procedure are shown in Table 2.5. Sodium bisulfite treated DNA was aliquoted and stored at -20ºC.

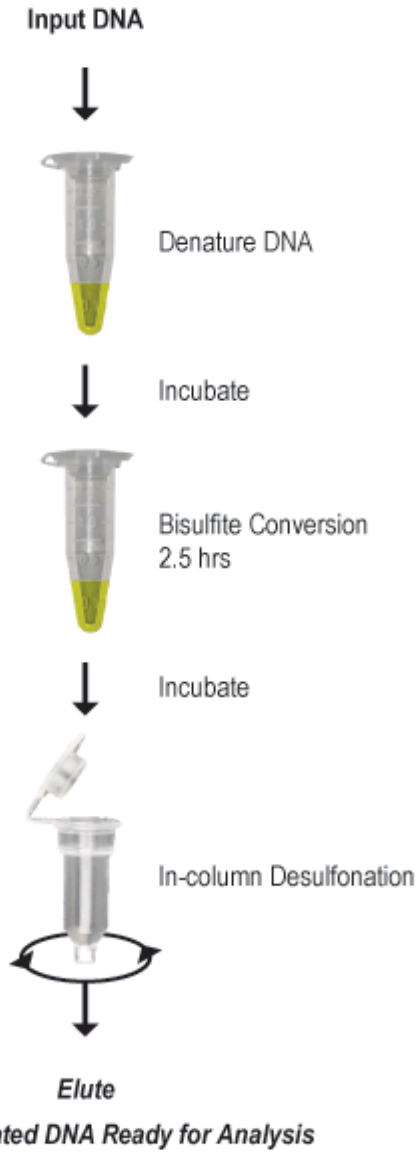


Figure 2.8 Overview of the sodium bisulfite treatment procedure

Step 1	**Add 130 µl of the CT Conversion Reagent to 20 µl of each DNA sample in a Conversion Plate. Mix the samples by pipetting up and down.**
Step 2	Seal the plate with the provided film. Transfer the **Conversion Plate** to a thermal cycler and perform the following steps: 1. 98°C for 10 minutes 2. 64°C for 2.5 hours 3. 4°C storage up to 20 hours
Step 3	Add 400 µl of **M-Binding Buffer** to the wells of **Silicon-A™ Binding Plate** mounted on a **Collection Plate**.
Step 4	Transfer the samples from the **Conversion Plate** (Step 2) to the wells of the **Silicon-A™ Binding Plate** and mix.
Step 5	Centrifuge at ≥ 3,000 x *g* (5,000 x *g* max.) for 5 minutes. Discard the flow-through.
Step 6	Add 400 µl of **M-Wash Buffer** to each well of the plate. Centrifuge at ≥ 3,000 x *g* for 5 minutes.
Step 7	Add 200 µl of **M-Desulfonation Buffer** to each well and allow the plate to stand at room temperature for 15-20 minutes. After the incubation, centrifuge at ≥ 3,000 x *g* for 5 minutes. Discard the flow-through.
Step 8	Add 400 µl of **M-Wash Buffer** to each well of the plate. Centrifuge at ≥ 3,000 x *g* for 5 minutes. Discard the flow-through. Add another 400 µl of **M-Wash Buffer** and centrifuge for 10 minutes.
Step 9	Place the **Silicon-A™ Binding Plate** onto an **Elution Plate**. Add 30 µl of **M-Elution Buffer** directly to each well. After 5 minutes, centrifuge at ≥ 3,000 x *g* for 3 minutes to elute the DNA.

Table 2.5 Bisulfite treatment steps. Adapted from the EZ-96 DNA Methylation-Gold Kit booklet (Zymo Research, Irvine, CA, USA).

2.4.2 Polymerase chain reaction

PCR was used to confirm the sodium bisulfite conversion. PCR is a technique commonly used to amplify a single or several copies of a piece of DNA, generating millions of copies of a particular DNA sequence. PCR uses a mix of components that are combined together and subject to cycles of heating and cooling in a thermocycler. During the first PCR step the mix is heated to a high temperature to activate the heat- sensitive polymerase taq. Three steps follow this:

. 1) A denaturation step, where the mix is heated to 95°C to denature the double-stranded DNA;

. 2) An annealing step where the mix is cooled to a primer-specific temperature (usually between 50°C and 65°C) to allow the primers to anneal with high specificity to the correct annealing sequence in the DNA;

. 3) An elongation step at 72°C to allow the taq polymerase to synthetize the complementary strand of DNA using the deoxynucleotides (dNTPs).

These three steps are repeated for a number of cycles to allow the synthesis of an exponential number of DNA amplicons. The final number of amplicons will be 2^n, where n is the number of cycles. A final step at 72°C is added to allow a final extension.

A PCR assay was performed to confirm sodium bisulfite conversion on a randomly selected sample of the genomic DNA (n=32), which includes both FF and FFPE tissues. Seven FFPE samples and 21 FF samples were randomly selected. The assay used primers designed to specifically target sodium bisulfite converted DNA in the CTCF binding site 3 region of the imprinting control region upstream of H19 (Chr11: 1979977 – 1980267). The forward primer used had the following sequence = 5' – TTTTAGGGTGTATTATTGAGGGTTA – 3', and reverse primer = 5' – CCCCATCCAAAAAAAACTTAAACTA – 3'.

The reaction components used in this thesis, their functions and volumes are detailed in Table 2.6. The heating and cooling cycles used are outlined in Table 2.7.

Component	Function	PCR reagent (concentration)	Quantity (µl)
Genomic DNA	Single-stranded bisulfite-treated* genomic DNA provides the reaction template	Bisulfite converted DNA (10 ng/µl)	1
Taq polymerase	Heat-resistant enzyme that extends primers to synthesise new strands of DNA complementary to the target sequence using DNA nucleotides	Qiagen HotStar Taq polymerase (5 units/µl) (Qiagen, Venlo, Holland)	0.05
DNA Primers	Short, single-stranded oligonucleotides complementary to the target DNA sequence	Forward and reverse primer mix (10µM)	1
DNA Nucleotides	Nucleotide bases required for the synthesis of the new DNA strands	dNTPs (2.5mM)	0.2
PCR buffer	Maintains optimum reaction pH	10 x reaction buffers	1
Magnesium chloride (MgCl2)	Required for Taq polymerase function	Magnesium chloride (25 mM)	0.2
RNAse/DNAse free water	Ensures consistent reaction volume	Water to 10 µl final volume	6.55

Table 2.6 Standard PCR reagents

Step	Function	Temperature (°C)	Time
1. Initiation	DNA mix is heated to activate the *Taq*. Hot-start *Taq* reduces mispriming and primer-dimer formation	95	15 minutes
2. Denaturation	High temperatures denature the double stranded DNA into single stranded DNA	95	20 seconds
3. Annealing	Lower temperatures allow the primers to anneal. Annealing temperature is selected carefully - too high and the primers are unable to anneal, too low and non-specific amplification occurs.	50 - 65	30 seconds
4. Extension	Taq polymerase uses dNTPs to build the complementary DNA strand	72	1 minute
Repeat steps 2 to 4 between 20 and 40 times		**Up to 40 repeats**	
5. Final extension	Remaining single stranded DNA is fully extended	72	3 minutes
6. Finish	Products are kept at low temperatures	4	∞

Table 2.7 Chosen PCR thermocycle.

2.4.3 Bisulfite DNA electrophoresis

PCR amplification was confirmed using agarose gel electrophoresis (Section 2.3.3) (Figure 2.9). There was a difference in general band intensity of bisulfite treated DNA between FFPE and FF samples on the gel electrophoresis images, with the FFPE samples generally lower in intensity. However, the presence of bands throughout indicated successful bisulfite conversion of genomic DNA.

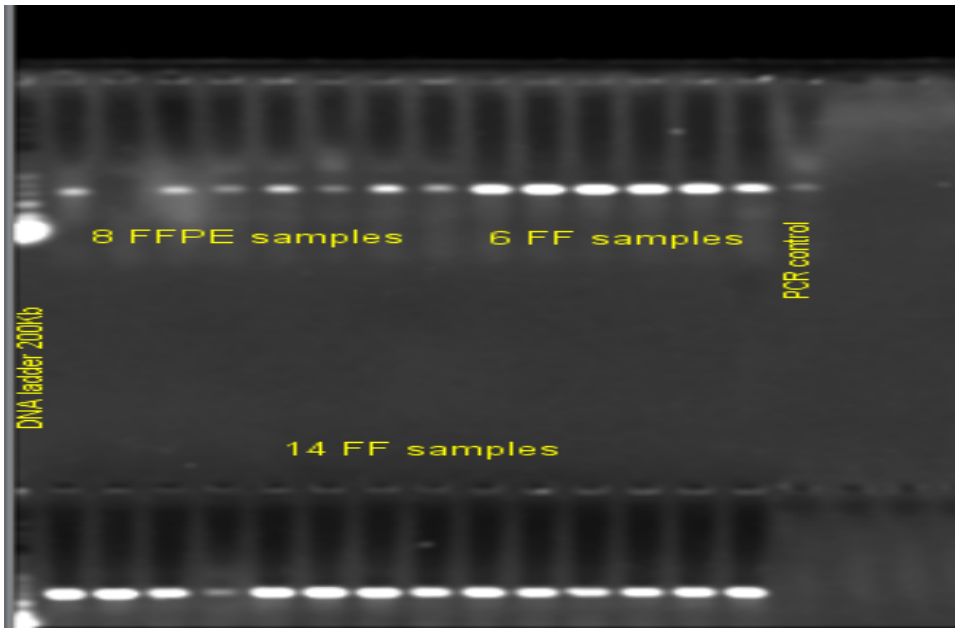


Figure 2.9 Agarose gel (2%) used to inspect PCR amplification products of random FF and FFPE samples. All samples amplified successfully during PCR except for FFPE sample 2 and FF sample 10. Top left shows 200bp ladder, 8 FFPE samples, 6 FF samples, water (negative control). Bottom left shows 200bp ladder, 14 FF samples. Whilst FFPE samples have amplified (except sample 2), it is noticeable that the intensity of FFPE samples is less than FF, indicating partially degraded/fragmented DNA in FFPE treated tissues.

2.4.4 QUBIT fluorometric quantitation

To ensure sufficient DNA concentration for downstream processing in the Illumina Infinium HumanMethylation450 Beadchip (Section 2.5), the concentration of genomic DNA in the samples was assessed with the QUBIT BR dsDNA assay using the QUBIT fluorometer (Life Technologies, Thermo Fisher Scientific). The QUBIT fluorometer is a small instrument that is used for

quantification of DNA, RNA and protein. It uses fluorescent dyes to quantify the concentration of nucleic acids and proteins in a sample. The dyes have low fluorescence until they bind to their targets, in this case the DNA. Once the dye binds to the DNA, it becomes intensely fluorescent. The fluorometer picks up the fluorescence signal and translates it into DNA concentration measurement. By using target-specific fluorescence, such as DNA, this method is more accurate than UV based methods for quantifying DNA concentration, such as the nanodrop. Table 2.8 illustrates the concentration of FF and FFPE samples randomly chosen.

Sample ID FF	Concentration (ng/ul)	Sample ID FF	Concentration (ng/ul)	Sample ID FFPE	Concentration (ng/ul)
FF1	2.35	FFPE9	1.48	FFPE1	2.37
FF2	2.31	FFPE10	1.29	FFPE2	2.25
FF3	1.3	FFPE11	3.4	FFPE3	0.981
FF4	9.3	FFPE12	4.86	FFPE4	1.16
FF5	4.77	FFPE13	0.513	FFPE5	1.52
FF6	4.66	FFPE14	4.51	FFPE6	1.71
FF7	7.53	FFPE15	1.82	FFPE7	0.85
FF8	5.52	FFPE16	3.56	FFPE8	0.879

Table 2.8 Concentration of FF and FFPE samples. As measured by the QUBIT fluorometer.

2.5 Genome-wide DNA methylation profiling

2.5.1 Introduction

Genome wide DNA methylation was quantified with the Illumina Infinium HumanMethylation450 BeadChip (Illumina, San Diego, CA, USA) with arrays run on the HiScan System (Illumina, San Diego, Cam USA). The samples were processed across 16 BeadChips. The Illumina HumanMethylation450 BeadChip

(450K array) profiles the DNA methylation status of 485,577 sites per sample. This method was selected for the quantification of DNA methylation status because it offers a range of advantages over alternative approaches; relatively low DNA input, simple sample preparation, high sample throughput, lower running costs and simpler data storage. Furthermore, the widespread use of this array within the epigenetics research community enables cross-study comparison of research findings. Finally, the extensive use of this technology has led to development of a wide variety of freely available bioinformatic tools that can be used to extend data analysis and aid interpretation.
Both FF and FFPE samples were examined using the Illumina HumanMethylation450 Beadchip. Due to the suboptimal quality of FFPE tissue and the chemistry, coverage and protocol of the 450K array, the final protocol was altered to input 7ul of the final eluate into the Infinium 450K array. The 450K arrays were then processed following the manufacturer's instructions.

The key features of the Illumina Infinium HumanMethylation450 Beadchip will be briefly outlined in the section below (Section 2.5.2).

2.5.2 Illumina HumanMethylation 450 BeadChip chemistry

The Illumina Infinium I Whole Genome Genotyping assay was originally developed for Single Nucleotide Polymorphism (SNP) genotyping (Steemers et al., 2006). The 450K array used in this project to quantify DNA is an adaptation of the Illumina Infinium I. The nucleotide variant produced by sodium bisulfite conversion of DNA (Section 2.4.1) can be quantitatively "genotyped" using Illumina technology (Bibikova et al., 2009).
The 450K array combines two assay types, Infinium I and Infinium II, to quantify DNA methylation status at single nucleotide resolution across the genome. The Infinium I assay was initially used by the predecessor of the 450K array, the Illumina HumanMethylation27 BeadChip (27K array). In this assay, two bead-bound probes per CpG locus, one "methylated" and one "unmethylated", are used to determine DNA methylation status (Figure 2.10). The 3' terminus of one probe targets the cytosine protected by the presence of DNA methylation

("methylated"), whereas the other probe targets the thymine that would result from bisulfite conversion and whole genome amplification of unmethylated DNA ("unmethylated"). DNA hybridisation is detected through a single-base extension with a labelled nucleotide, fluorescent staining and scanning. Each probe is 50 base pairs (bp) in length, therefore the design for Infinium I assays assumes that methylated is regionally correlated within 50 bp span. This assumption of regional correlation is supported by a study that found over 90% of CpG sites within 50 bases had the same DNA methylation status (Eckhardt et al., 2006), and the finding that methylation at adjacent sites tends to be correlated (Shoemaker et al., 2010).

The Infinium II assay was introduced on the 450K array. This assay uses one probe per CpG locus rather than two, which enables increased genomic coverage, as more beads can be included on the array. In this assay, the 3' terminus of the 50 bp probe complements the base directly upstream of the query site. The single base extension results in the incorporation of either a labelled G or A base, complementary to either the 'methylated' C or "unmethylated" T. The methylated and unmethylated signals are distinguished through use of different dye colours (green and red, respectively) (Figure 2.10). Additionally, the Infinium II probes are designed without the assumption that nearby CpGs are methylated to the same extent as the target CpG. Instead, degenerate bases ("R") are included at CpG sites to allow assessment of methylation status at the query site independently of other CpG sites. However, a maximum of three R bases can be included in a probe sequence without data compromise, therefore the 450K array also includes a number of Infinium I probes to allow assessment of regions of DNA with high CpG density, such as promoter CpG islands.

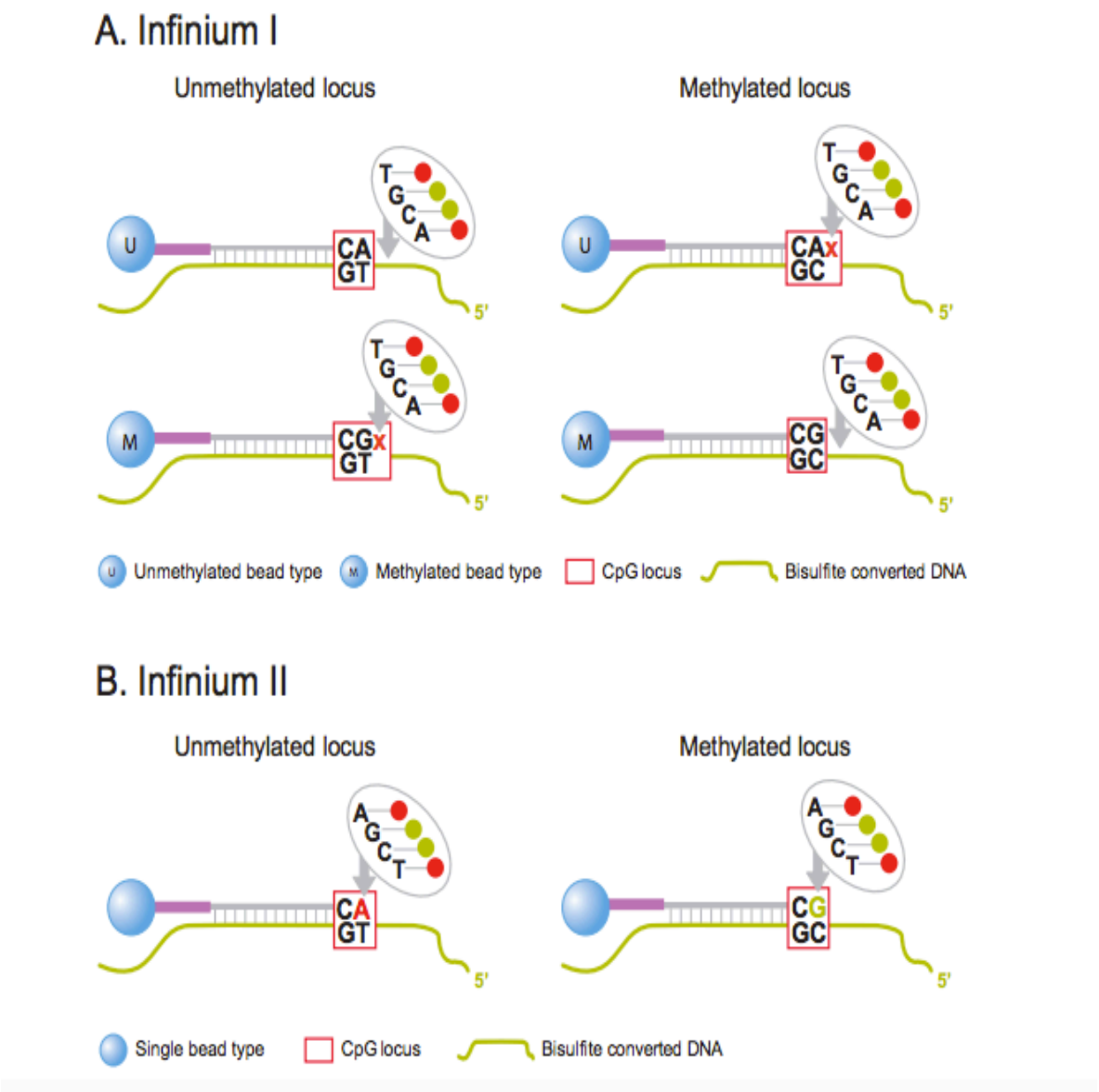


Figure 2.10 Infinium I and Infinium II chemistry. A) Infinium I employs two bead types per CpG locus, one 'methylated' and one 'unmethylated' B) Infinium II design uses one bead per locus. The methylation state is determined at the single base extension step after hybridisation.

2.5.3 Genomic coverage of the Illumina HumanMethylation450 BeadChip

The Illumina 450K BeadChip combines Infinium I and Infinium II assay chemistries for greater depth of coverage, enabling DNA methylation quantification at single nucleotide resolution at 485,577 sites across the

genome (n = 135,501 Infinium I; n = 350,076 Infinium II) (Section 2.5.2). By allowing the selection of specific target sequences through probe design, this technology avoids the bias that is inherent in capture methods, which typically miss regions of low CpG density, for example. However, the array only allows the investigation of a pre-designed probe set, which introduces the challenge of selecting appropriate CpG sites for inclusion on the array. A consortium of epigenetics researchers guided the selection of CpG sites to be interrogated by the 450K array (Bibikova et al., 2011). Emphasis was placed on covering CpG islands and RefSeq genes. Consequently 96% of CpG islands are interrogated and an average of 17 CpG sites across the 5'UTR, first exon, gene body and 3'UTR of 99% of RefSeq genes (Bibikova et al., 2011). Additional sites covered by the array include CpG sites outside CpG islands, non-CpG sites identified in human stem cells, miRNA promoter regions and approximately 90% of all sites covered by the Illumina 27K array (482,421 CpG sites, 3091 non-CpG sites, 65 control SNPs).

2.6 Statistical analysis

2.6.1 Quality control pipeline

2.6.1.1 DNA data pre-processing

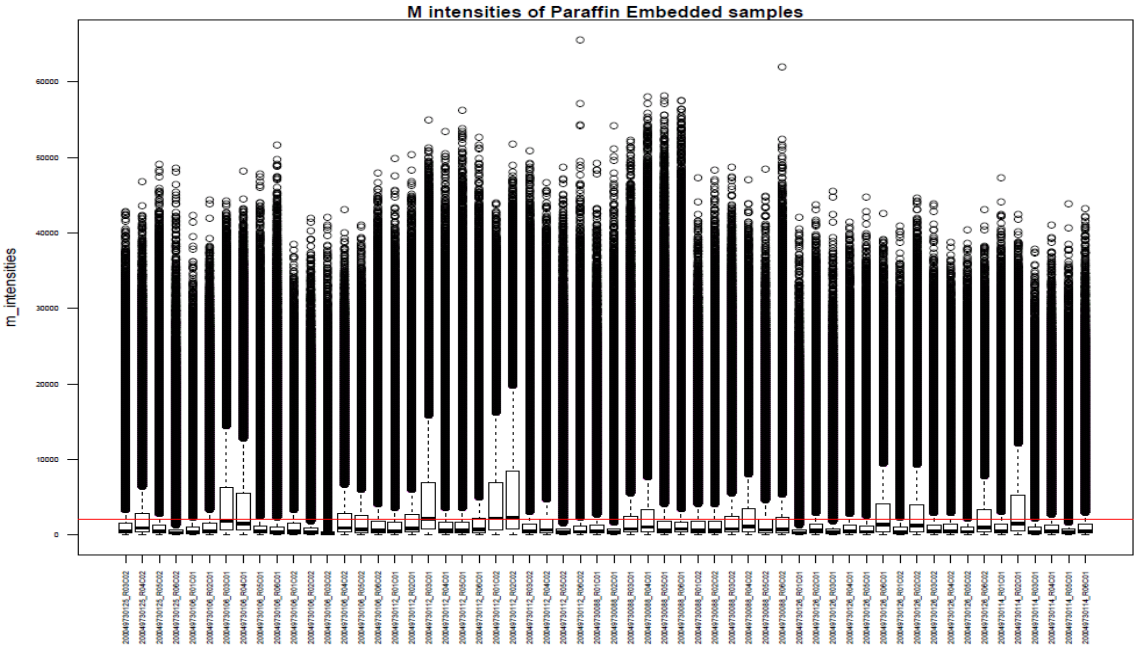
Following the genome wide DNA methylation using the Illumina 450K BeadChip Array, signal intensities for each probe were extracted using the Illumina GenomeStudio Software (Illumina, San Diego, CA, USA). Data from 45 unique individuals (90 matched samples) entered our Quality Control (QC) pipeline, which was performed in the R (www.r-project.org) statistical programming environment (www.bioconductor.org/packages/release/bioc/html/methylumi.html) package. The *wateRmelon* package (Pidsley et al, 2013) was used for quality control and pre-processing of the data as described below.

2.6.1.2 Sample and probe removal

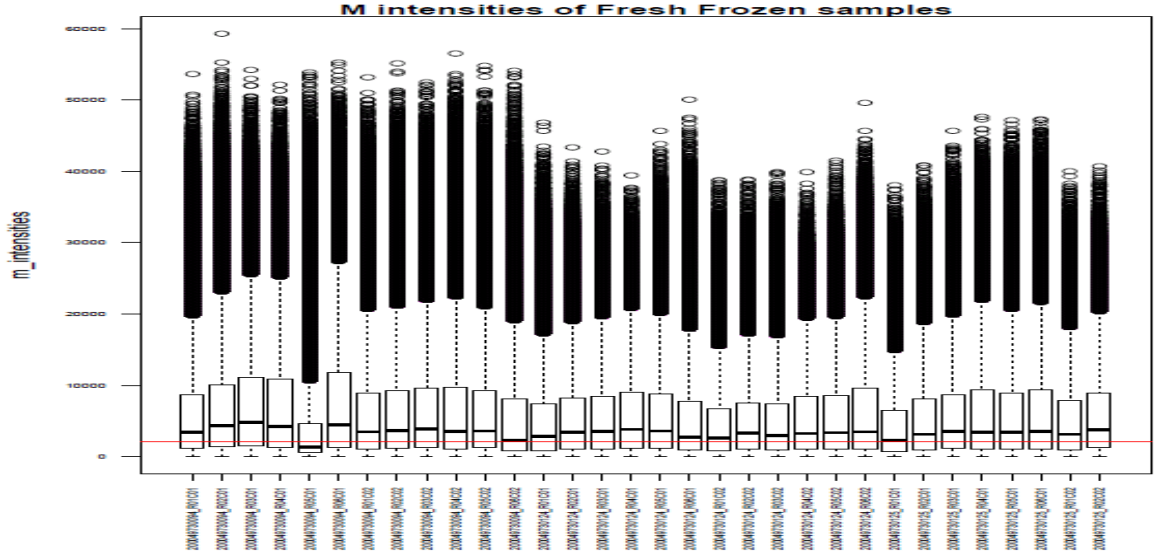
DNA Methylation levels reported as β values, are calculated from mean methylated (m) and unmethylated (u) signal intensities for each locus for each

sample. Therefore, the first step in the QC pipeline was to stringently filter the unnormalised Illumina 450K data using the *pfilter* function from the *wateRmelon package* (Pidsley et al, 2013). This function was used to filter CpG loci and samples by bead count and detection P-Values. This assessed the signal intensities of the samples. There was a marked difference in signal intensities between FF and FFPE tissues. FF samples demonstrated higher signal intensities than FFPE samples (Figures 2.11). Then the efficiency of bisulfite conversion was examined to exclude samples with conversion score <80. Four FFPE samples were found to have a conversion score <80.

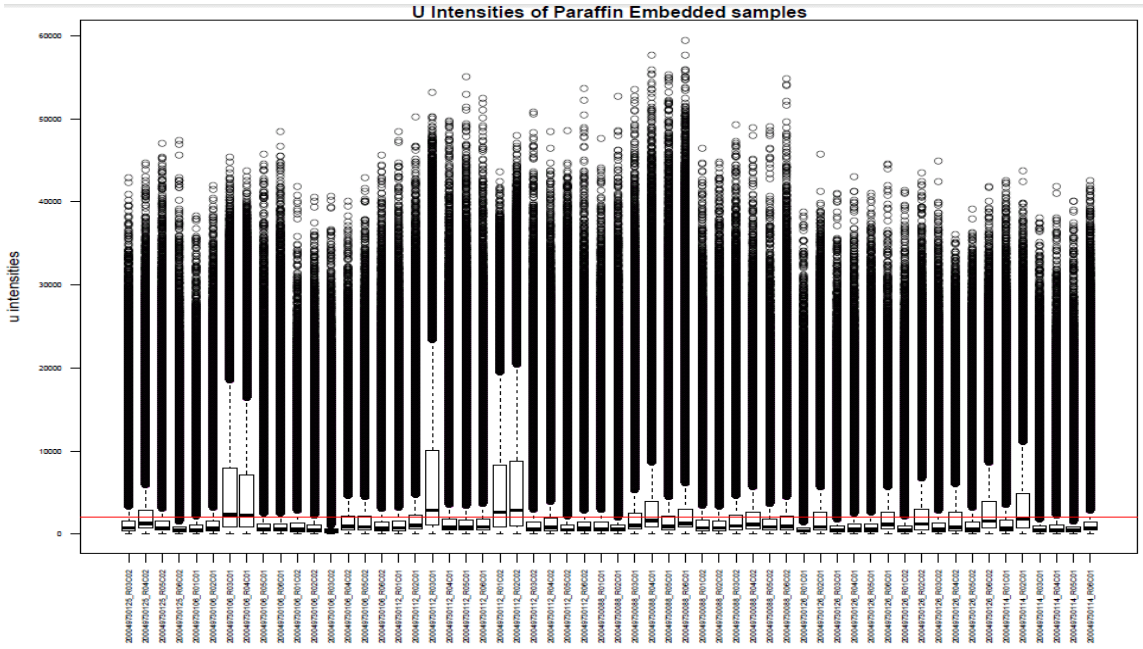
A



B



C



D

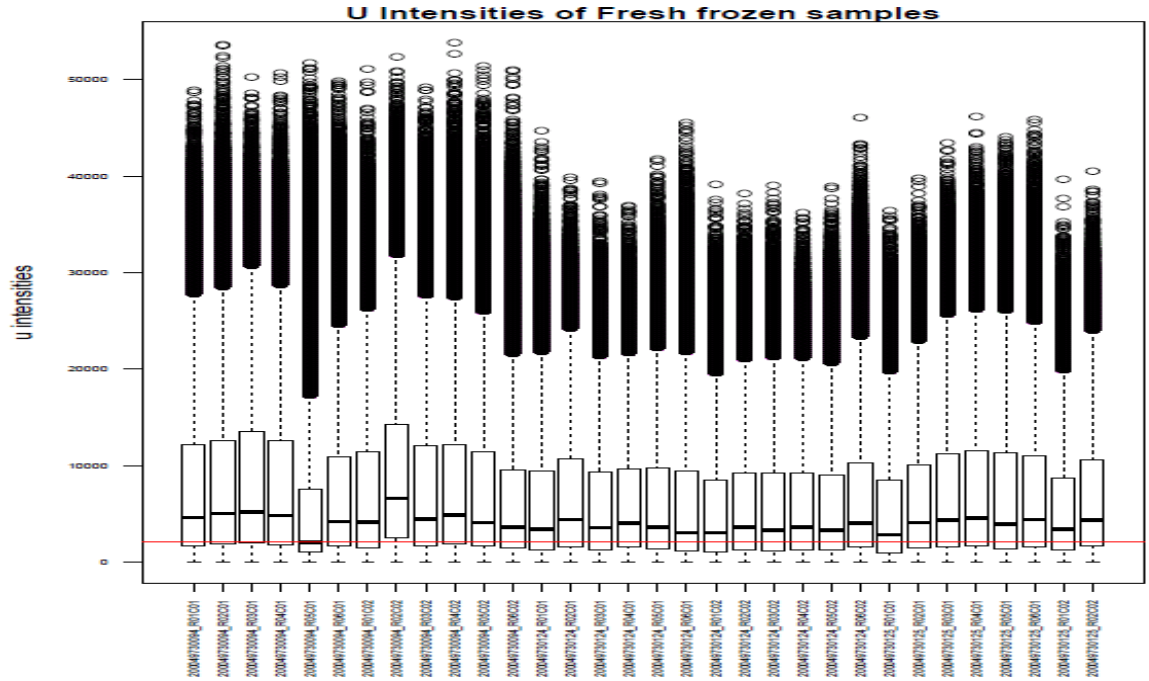


Figure 2.11 Methylated and unmethylated signal intensities across all samples. A) shows the methylated intensities of the FFPE samples, with the red line representing the cut off threshold of 2000. B) In contrast to A, the majority of the FFPE sample intensities in B were below the threshold. The same applies to the unmethylated intensities of FFPE samples in C) vs unmethylated intensities of FF samples in D).

2.6.1.3 Principle component analysis

To visually assess similarities and differences between the two types of samples and determine whether they can be grouped, principle component analysis (PCA) algorithm was applied to the samples. PCA executed on the whole data set showed a clustering of samples based on treatment method (FF and FFPE) (Figure 2.12). A second PCA of all samples against median methylation intensities showed a clustering of the FF samples (with the exception of one sample) above 2000, and a clustering of FFPE samples (with the exception of three samples) <2000 (Figure 2.13).

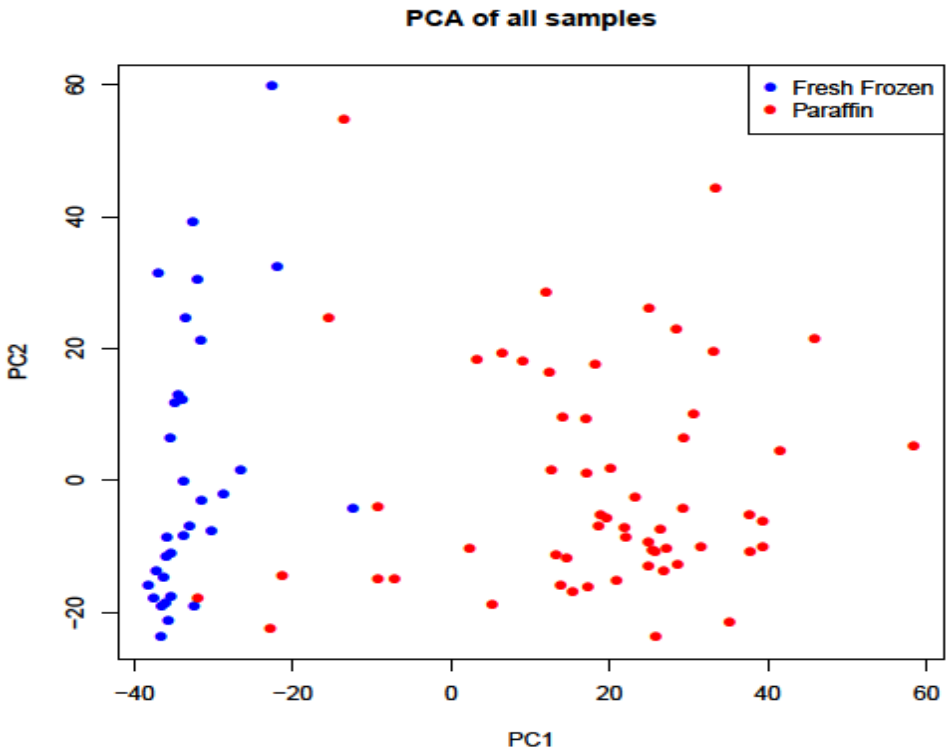


Figure 2.12 PCA of all samples. This shows the clustering of samples based on treatment method: the FF samples are clustered in blue to the left of the graph, whereas the FFPE samples are clustered to the right of the graph in red.

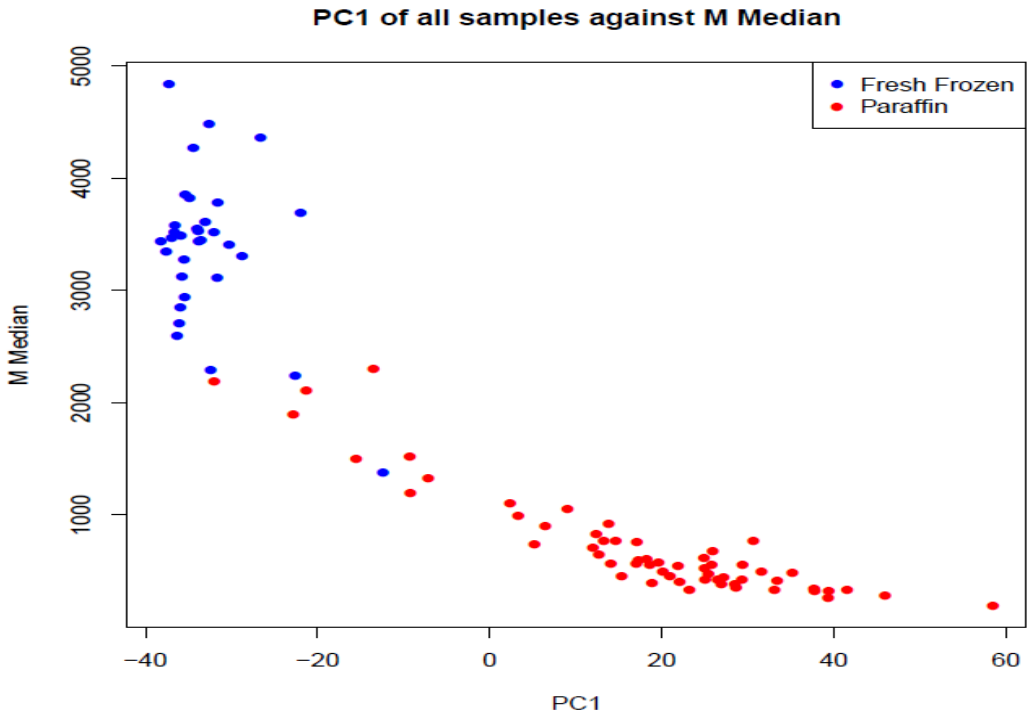


Figure 2.13 PCA of all samples against M median. This shows a clustering of most of the FF samples (except one) above the 2000 threshold of median methylation intensity (in blue). The majority of FFPE samples are clustered in red below 2000.

2.6.1.4 Selecting samples that passed QC

Following the QC checks, in the light of the low signal intensities of the FFPE samples and the clustering of samples using PCA graphs, samples with both the median methylated 'M' and unmethylated 'U' intensities <2000 and their matched counterparts (matched cancer or normal mucosal) were excluded from subsequent analysis. Therefore, all FFPE samples and two FF samples were excluded.

A final dataset comprising of n = 30 samples remained following implementation of QC procedures. These were all FF matched rectal tumour and adjacent normal mucosa samples from 15 individuals. Therefore, the term FF in the rest of this section will refer to the FF samples that passed our stringent QC methods and the rest of the data analysis was only performed on those samples.

2.6.1.5 Data normalisation

Following QC checks, normalisation of data is necessary to remove systematic differences between samples and reduce undesirable variation between arrays. This ensures maximum sensitivity to detect differential methylation between experimental groups. Here, data were normalised using quantile-normalisation (QN). QN is a non-linear transformation that produces identical array-wide distributions by replacing each intensity score with the mean of the features with the same rank from each array. This is a well-established approach within gene expression analysis, where it has been found to perform well (Irizarry et al., 2003). Specifically, the "*dasen*" method was utilised for data normalisation. This approach, defined by Pidsley *et al*., involves background adjustment of methylated and unmethylated intensities, followed by separate QN of methylated Type I, unmethylated Type I, methylated Type II and unmethylated Type II intensities. Using three performance metrics (utilising known DNA methylation patterns associated with genomic imprinting, X-chromosome inactivation and SNP genotyping assays included on the array) this normalisation scheme was quantitatively determined to be optimal (Pidsley et al., 2013).

2.6.1.6 Gender check

The cohort was assessed using multi-dimensional scaling (MDS) to graphically visualise the similarity of individuals within a dataset. The concordance between the expected sample sex (collected data), and the reported sample sex (DNA methylation data) was assessed through MDS visualisation of the DNA methylation status of Y-chromosome and X-chromosome CpG sites, which are expected to display sex differences due to the process of X-chromosome inactivation in females. There were no discrepancies in the reported gender and sex chromosomes in any of the samples that passed QC (Figure 2.14).

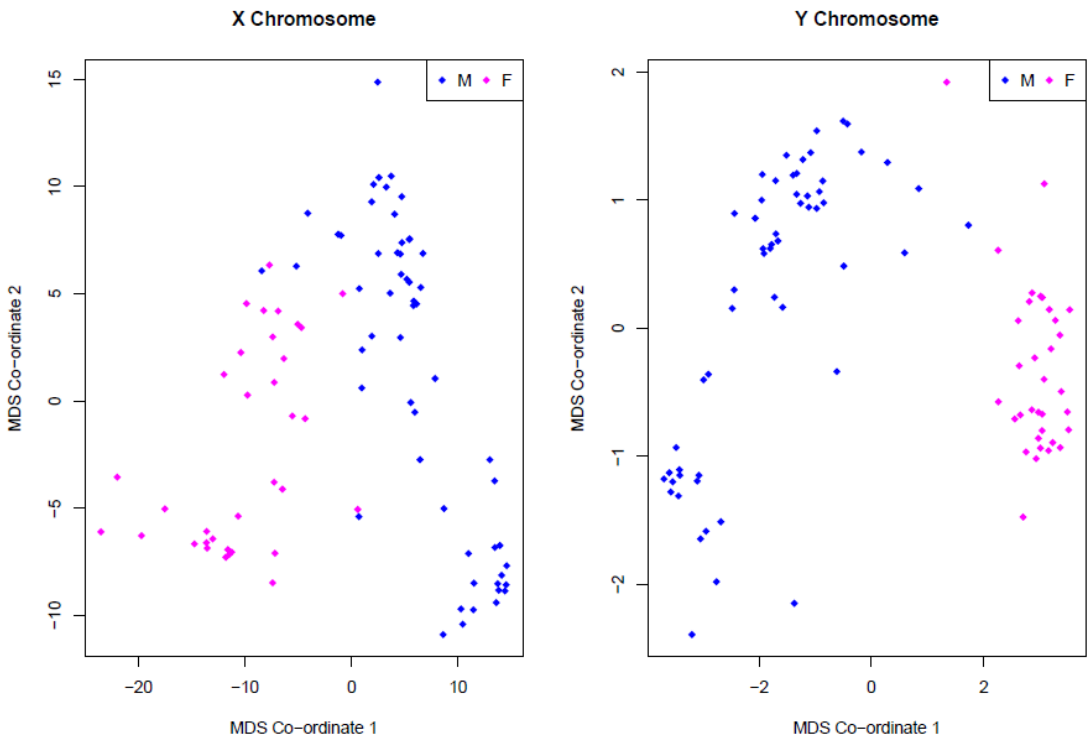
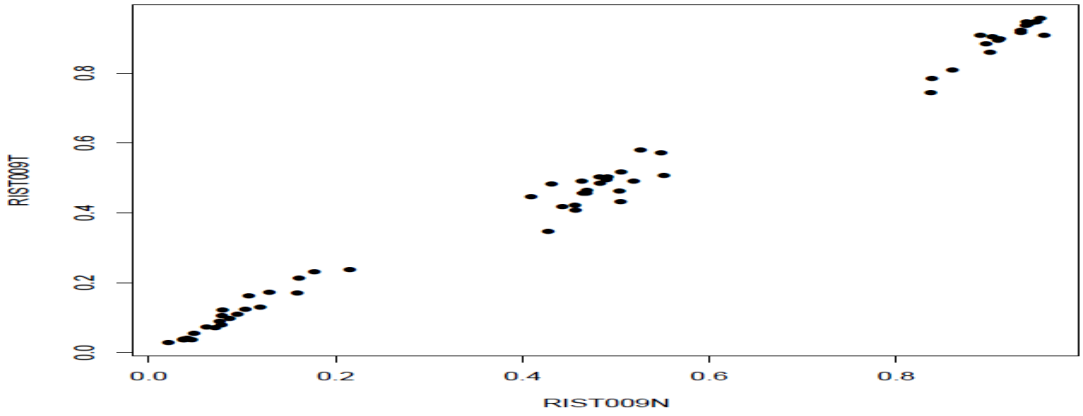


Figure 2.14 Multi-dimensional scaling for gender check. No discrepancies between reported and detected gender were found.

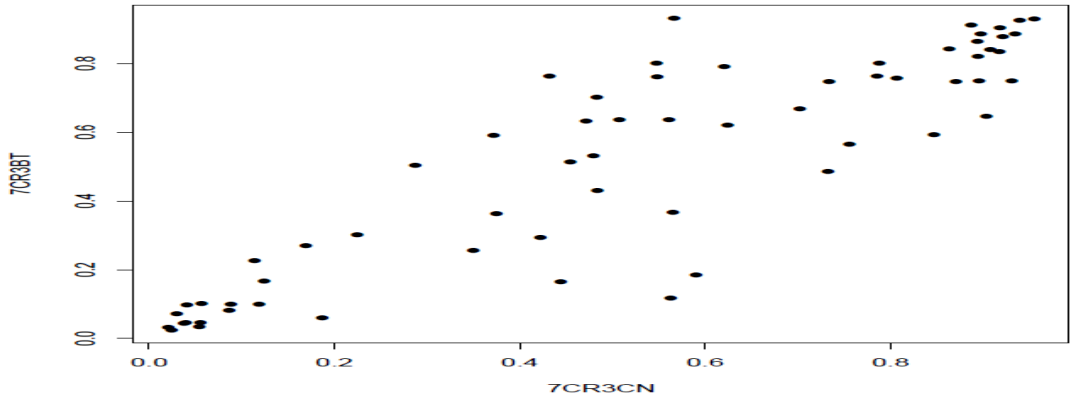
2.6.1.7 Sample identity check

In addition to examining predicted and reported sex, sample identity was further validated by examining the genotype correlation between the two types of tissues: rectal tumour and normal mucosa. There are 65 SNPs included on the 450k methylation array that can be used to generate a DNA "fingerprint" of samples as an added level of QC. This is done by plotting the beta values from the 65 SNPs methylation profile in a scatter plot. Samples from the same individual have the SNP results fall along the identity line in a scatter plot whereas samples from different persons scatter into the 9 different possible spots, based on their genotypes (Figure 2.15). SNP results of the samples that did not pass the QC would not be expected to pass the sample identity validation.

A



B



C

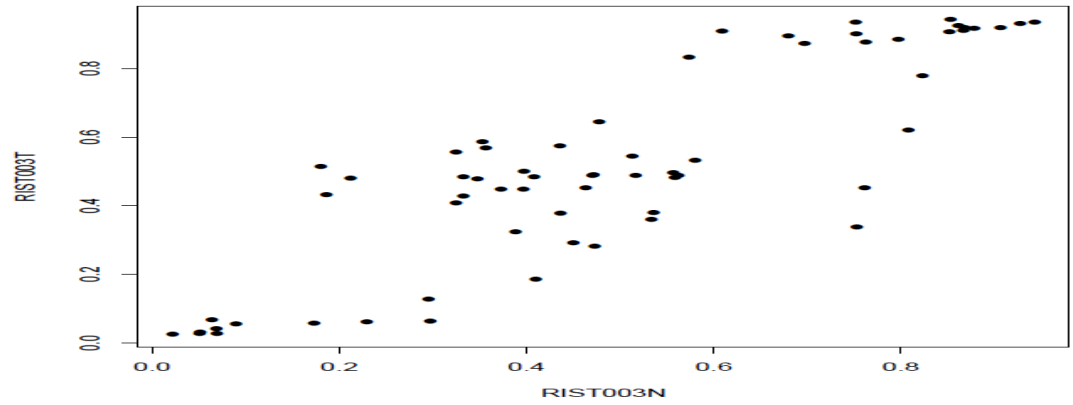


Figure 2.15 Scatter plots of SNP results from three samples. A) The graph shows good correlation with SNP results from the same individual falling along the identity line. B) and C) show a scattering of the results with dots more spread out across the nine quadrants. B) represents a FFPE sample. C) represents a FF sample – the only FF sample to fail the QC test.

2.6.1.8 DNA methylation age calculation

DNA methylation values for the FF samples were uploaded to the DNA methylation online age calculator (Hovrath, 2013) to estimate DNA methylation age for each sample and obtain tissue prediction information. The DNA age calculator was developed using 8,000 samples from a broad range of healthy tissues, cancer tissues and cell lines. The calculator allows the estimation of DNA methylation age from a complete 450k dataset, using a weighted average data from 353 'clock CpGs', which is then transformed to DNA methylation age using a calibration function. DNA methylation age for each FF sample is presented in Supplementary Table 1 (Appendix 4).

2.6.2 Data analysis

2.6.2.1 Paired t-test

Following successful QC, data was analysed using a t-test for group mean differences in DNA methylation between the normal and tumour samples. Global methylation difference was reported.

For each CpG site a β-value was generated by calculating the ratio of the normalised signal from the methylated probe to the sum of the normalised signals of the methylated and unmethylated probes. Consequently, the β-value is a measurement ranging continuously from zero (unmethylated) to one (fully methylated).

2.6.2.2 Top probes selection

To identify rectal cancer associated DNA methylation differences at the individual probe level across all samples, I performed a linear regression using the pre-processed and normalised DNA methylation β values as the dependent variable and tissue type, cancer vs normal, as the independent variable. Other phenotypical factors such as medication, smoking status and gender were not included due to incomplete data.

DNA methylation differences are frequently auto-correlated across the genome, therefore probe-wise multiple testing corrections are often overly stringent, which can make data-interpretation challenging. To account for concordance of

DNA methylation effect across genomic regions, the command line tool *comb-p* was used to identify spatially correlated DNA methylation-phenotype associations. Illumina probe-site chromosomal coordinates and respective P-values for the effects of interest were output from the linear regression performed as described above. This data was then manipulated using the program *comb-p* to (1) calculate auto-correlations (2) combine adjacent P values (3) perform false discovery adjustment (4) find regions of enrichment and (5) assign significance to those regions.

2.6.2.3 Top regions selection

Because DNA methylation is often correlated across adjacent CpG sites across broad genomic regions we also sought to identify differentially methylated regions (DMRs) associated with rectal cancer in our samples. Although several approaches exist to identify DMRs, we used the Python module, *Comb-P,* to detect spatially correlated P-values (Pedersen et al, 2012) and group spatially correlated differentially methylated probes (seed P-value < 1.00E-3, minimum of 5 probes) in each region. The regions' p-values were corrected for multiple testing using Sidak correction (Sidak, 1967), which corrects the combined P for na/nr test, where na is the total number of probes tested in the initial EWAS and nr the number of probes in the given region.

2.6.2.4 Gene ontology analysis

An enrichment analysis on Gene Ontology (GO) terms was performed next, to test for a significant enrichment of particular GO terms (Ashburner et al., 2000). A logistic regression model developed by our group was used, as described in Lunnon et al. (2016), to test if genes in each module predicted pathway membership, while controlling for the number of probes in each network annotated to each gene. Pathways were downloaded from the GO website (Gene Ontology Consortium, 2015) and mapped to genes including all parent ontology terms. All genes with at least one 450k probe annotated and mapped to at least one GO pathway were considered. Pathways were filtered to those containing between 10 and 2000 genes. After applying this method to all pathways, the list of significant pathways (P < 0.05) was refined by grouping

related pathways to control for the effect of overlapping genes. This was achieved by taking the most significant pathway, and retesting all remaining significant pathways while controlling additionally for this best term. If the test genes no longer predicted the pathway, the term was said to be explained by the more significant pathway, and hence these pathways were grouped together. This algorithm was repeated, taking the next most significant term, until all pathways were considered as the most significant or found to be explained by a more significant term.

Chapter 3 - Results

3.1 Methodological overview

Matched tumour and normal adjacent mucosal colorectal samples were collected and DNA isolated using methods described (Section II). Genomic DNA was treated with sodium bisulfite in duplicate using the EZ-96 DNA methylation-gold kit (Zymo Research, Irvine, CA, USA) and DNA methylation profiled using the Infinium HumanMethylation450 BeadChip (Illumina, San Diego, CA, USA) processed on an Illumina HiScan System (Illumina) using an adjusted protocol (Section 2.5.1). All samples were randomised within and between arrays to avoid potential batch effects. In total, 90 FF and FFPE matched rectal samples, of normal mucosa and adjacent tumour, entered the QC pipeline. The majority (> 90%) of the FFPE samples and one FF sample failed the QC test. Therefore, all FFPE samples (n= 62) and two FF samples (the tumour sample that failed the QC test and its matched normal sample counterpart) were excluded. In total, 30 matched samples, all FF, passed the stringent QC tests. The data from these samples was subsequently analysed. In summary, we assessed genome wide patterns of DNA methylation in 30 human matched rectal tumour and adjacent normal mucosal samples (7 males and 8 females, age range 45-79 years). The results obtained from the QC analysis of the FFPE samples will be presented first, followed by the results of the FF samples and the methylation data obtained.

3.2 FFPE tissues

3.2.1 Clinical and histopathological characteristics

Historical CRC samples were collected from 2004-2007 and stored in the ETB database. Sections of these samples were cut, placed in micro-cassettes, immediately fixed in formalin and treated with paraffin. Paraffin embedding allows for histological verification of the content and quality of the historical samples, and assesses the integrity and heterogeneity of the tumour cells. All historical samples were thought to be adenocarcinoma tumours taken from the colonic epithelium. However, the results of the histological examination showed not all adjacent mucosal samples were biopsied from the colonic epithelium and tumour samples contained varied cancer content (visually estimated in percentage terms). Table 3.1 shows a summary of the histological findings.

Sample	Type of tissue collected adjacent to the tumour	Cancer % in rectal tumour tissue collected
1	Colonic epithelium	20%
2	Colonic epithelium	100%
3	Colonic epithelium	90%
4	Muscle sample	35%
5	Ischemic sample	0% (Ischemic sample)
6	Dysplastic sample	100%
7	Colonic epithelium	90%
8	Colonic epithelium	30%
9	Colonic epithelium	10% Tumour + Necrosis
10	Colonic epithelium	50% Tumour + Necrosis
11	Colonic epithelium	90%
12	Colonic epithelium	55%
13	Colonic epithelium	90%
14	Dysplastic sample	50%
15	Muscle Sample	20%
16	Colonic epithelium	45%
17	Colonic epithelium	90%
18	Colonic epithelium	0% (Dysplastic sample)
19	Colonic epithelium	90%
20	Colonic epithelium	70%
21	Colonic epithelium	30%
22	Colonic epithelium	80%
23	Colonic epithelium	80%
24	Colonic epithelium	90%
25	Colonic epithelium	35%
26	Colonic epithelium	40%
27	Colonic epithelium	50%
28	Colonic epithelium	0% (Dysplastic sample)
29	Colonic epithelium	0% (Dysplastic sample)

Table 3.1 Type of tissue in FFPE mucosal samples and quantification of cancer in tumour samples.

3.2.2 QC and exclusion from methylation analysis

Before proceeding with the Illumina 450K microarray protocol, it was necessary to sodium bisulfite convert all the samples (n=90). Genomic DNA (500ng) from each individual was treated with sodium bisulfite using the Zymo EZ DNA MethGold Kit according to the manufacturer's protocol **(**Section 2.4.1).

The first QC step assessed the methylation intensity profiles of all the samples. Prior to methylation quantification using Illumina 450K, a PCR assay on a portion of the samples was performed to confirm the success of the sodium bisulfite conversion. Seven FFPE samples and 21 FF samples were randomly selected. PCR amplification was confirmed using agarose gel electrophoresis. There was a difference in general band intensity of bisulfite treated DNA between FFPE and FF samples on the gel electrophoresis images, with the FFPE samples generally lower in intensity. The presence of clear bands throughout however, indicated successful bisulfite conversion of genomic DNA (Section 2.4.3) (Figure 2.9). To ensure sufficient DNA concentration for downstream processing in the Illumina Infinium HumanMethylation450 Bead Chip, the concentration of genomic DNA in the samples was assessed with the QUBIT BR dsDNA assay using the QUBIT Fluorometer (Life Technologies, Thermo Fisher Scientific). This method is more accurate than UV based methods such as the nanodrop, for quantifying DNA concentration. QUBIT Fluorometric Quantification showed that the DNA concentration of FFPE samples was overall lower than the FF samples (DNA concentration mean for FF samples = 4.71, DNA concentration mean for FFPE samples = 2.02), yet sufficient to proceed with downstream processing (Section 2.4.4). Both FF and FFPE samples were examined using the Illumina HumanMethylation450 BeadChip. Due to the suboptimal quality of FFPE tissue and their slightly lower DNA concentration, and due to the chemistry, coverage and protocol of the 450K array, the final protocol was altered to input 7ul of the final FFPE elute, more than the standard protocol, into the Infinium 450K array. The 450K arrays were then processed following the manufacturer’s instructions.

Following the genome wide DNA methylation using the Illumina 450K BeadChip Array, signal intensities for each probe were extracted using the Illumina GenomeStudio Software (Illumina, San Diego, CA, USA). Data from 45 unique individuals (90 matched samples) entered our Quality Control (QC) pipeline, which was performed in the R (www.r-project.org) statistical programming

environment. QC tests returned suboptimal results on the FFPE samples. The first QC step was to measure the signal intensities of the samples. DNA Methylation levels reported as β-values, are calculated from mean methylated (m) and unmethylated (u) signal intensities for each locus for each sample. There was a marked difference in signal intensities between FF and FFPE tissues. FF samples demonstrated higher signal intensities than FFPE samples (Figure 2.11). In addition to examining predicted and reported sex, sample identity validation is another important QC measure. Sample identity was validated by examining the genotype correlation between the two types of tissues: rectal tumour and normal mucosa. There are 65 SNPs included on the 450K methylation array that can be used to generate a DNA "fingerprint" of samples as an added level of quality control. This is done by plotting the β-values from the 65 SNPs methylation profile in a scatter plot. Samples from the same individual have the SNP results fall along the identity line in a scatter plot, samples from different persons scatter into the 9 different possible spots, based on their genotypes (Figure 2.15). SNP results of the samples that did not pass the QC would not be expected to pass the sample identity validation.
A high correlation between the two types was observed for all FF samples (except one sample) further confirming no sample mix-ups and success of detection of methylation intensities. However, a low correlation ($r<0.95$) between rectal tumour and normal mucosal tissue was found in all FFPE samples (Section 2.6.1.7). Finally, PCA executed on the whole data set showed a clustering of samples based on treatment method (FF and FFPE). A second PCA of all samples against median methylation intensities showed a clustering of the FF samples (except for one sample) above 2000, and a clustering of FFPE samples (except for three samples) <2000 (Figure 3.1). Following these results, the threshold for passing the QC check was set at median methylation intensity of >2000 for methylated (m) and unmethylated (u) signals. This threshold was set based on previously tested thresholds within the group, the QC tests showing suboptimal results for the FFPE samples and a further PCA of all samples against methylation intensities, showing a clear separation of the two types of tissues above and below the 2000 mark. Only three FFPE samples (out of 58 analysed) passed the threshold. Figure 2.11 shows the signal intensities of all FFPE samples. Consequently, all FFPE samples were excluded from further analyses. All FF samples except for one had a

methylation signal intensity of >2000. After exclusion of the one FF sample with signal intensity <2000 and its matched counterpart, 30 samples in total all processed as FF were included in the methylation analyses.

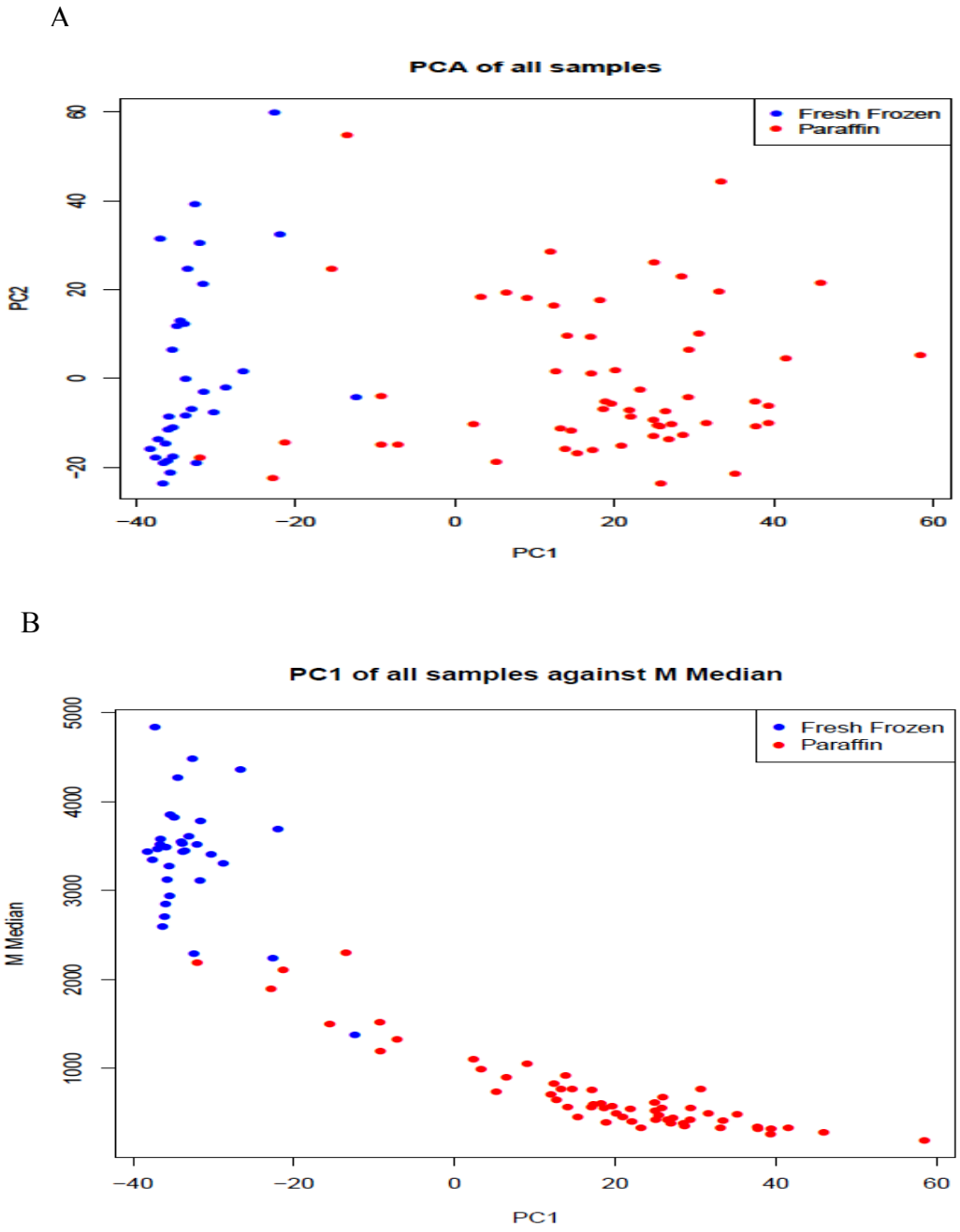


Figure 3.1 PCA showing clustering of samples. A) PCA executed on the whole data set showed a clustering of samples based on treatment method (FF and FFPE). B) PCA of all samples against median methylation intensities showed a clustering of the FF samples (with the exception of one sample) above 2000, and a clustering of FFPE samples (with the exception of three samples) <2000.

3.3 FF samples

3.3.1 Introduction

Following a successful DNA extraction and bisulfite conversion, the samples were processed using Illumina 450K. The unnormalised Illumina 450K data was stringently filtered using the *pfilter* function of the *watermelon* package. This, filters both samples and CpG loci dependent on bead count and detection p-values. Following data normalisation, PCA was applied to reveal the presence of any underlying data-structure that could be attributed to unknown variables. PCA executed on the FF samples included in the analysis set showed a clustering of samples based on pathological disease state (Tumour vs Normal), indicating the presence of widespread epigenetic differences associated with rectal cancer (Figure 3.2). Following successful QC, data was analysed using a t-test for group mean differences in DNA methylation between the normal and tumour samples. No further covariates were included in this test. A linear regression model was used to identify DNA methylation at individual probes then region-level analysis was performed by spatially combining correlated P-values using the *Python* module *comb-p*. Finally, Gene Ontology (GO) enrichment analysis was performed.

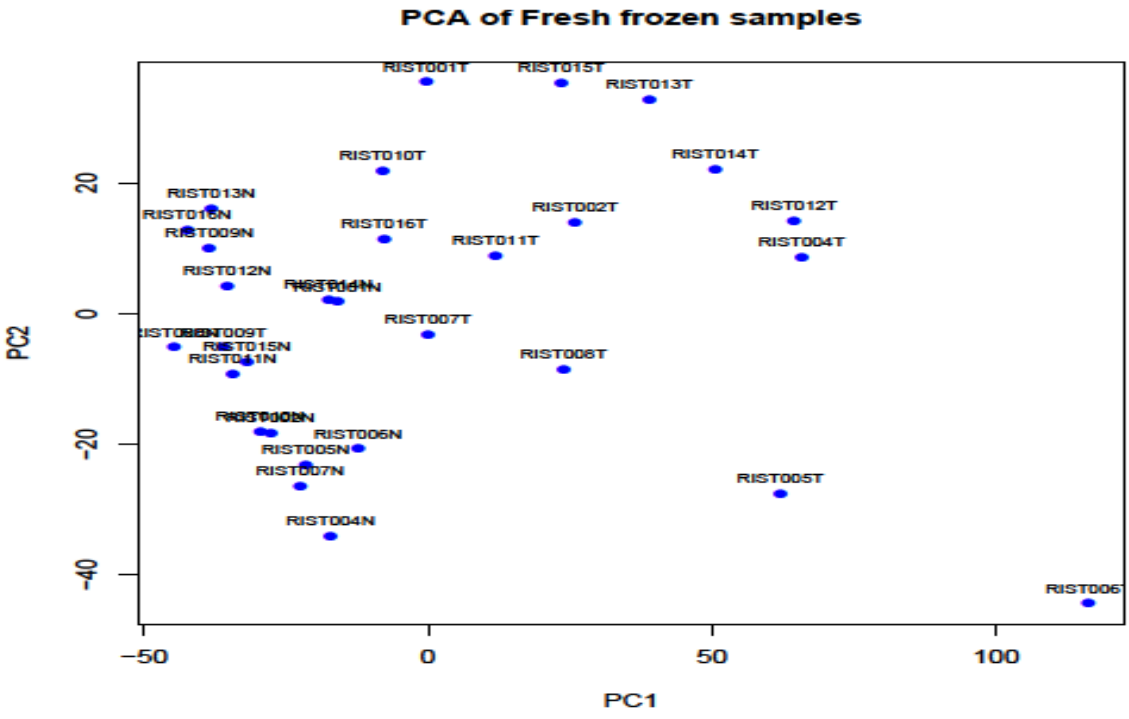


Figure 3.2 PCA of FF samples. PCA on all FF samples showed a clustering of samples based on type of tissue (Normal mucosa [RISTN] Vs Tumour [RISTT]).

3.3.2 Clinical and pathological characteristics of rectal cancer patients

A final data set comprising of n=30 FF samples and n = 424,413 probes remained following implementation of QC procedures. These samples corresponded to rectal cancer patients and included n=7 males (age range 45-79) and n=8 females (age range 63-79). The clinical and pathological characteristics are described in Table 3.2 and Table 3.3.

We analysed rectal tumour tissue and matched normal adjacent mucosal tissue from 15 patients diagnosed with rectal tumours (adenocarcinoma n= 12, high grade dysplasia n= 2, low grade dysplasia n=1) at the Royal Devon and Exeter hospital Exeter, Devon, UK.

The inclusion criteria were as follows: a clinical diagnosis of rectal cancer (histopathological initial polyp biopsies were considered unrepresentative due to their superficial nature. Subsequently, three samples showed dysplasia, and the rest (n=12) were adenocarcinoma). The median age of the patients was 71 years (range 44.9 – 74.7 years) with 7 men and 8 women.

Samples were taken during surgery or colonoscopy, after the purpose and nature of all the procedures were fully explained and written consent obtained from all patients. Biopsies from all 15 patients contained a sufficient amount of tissue for DNA extraction and downstream processes. The CRF NIHR in Exeter, approved the protocol of the study.

Characteristics	**N**	**%**
Gender		
Female	8	53.33
Male	7	46.67
Age		
Median	71	
Range	45-79	

Procedure		
Surgery	15	100.00
Type		
Adenocarcinoma	12	80.00
High Grade Dysplasia	2	13.33
Low Grade Dysplasia	1	6.67
Tumour (Radiological Stage)		
T4	2	13.33
T3	8	53.33
T2/3	3	20.00
T2	1	6.67
N/R	1	6.67
Nodes (Radiological Stage)		
N2	1	6.67
N1	5	33.33
N0	8	53.33
N/R	1	6.67
Metastasis		
No	15	100.00
Yes	0	0.00
Location		
High Rectum	8	53.33
Mid Rectum	1	6.67
Low Rectum	6	40.00

Table 3.2 Summary of the clinical and histopathological characteristics of tumour samples.

Sample	Age	Gender	Ethnicity	Tumour Location	Histology	Rad T	Rad N	Hist T	Hist N	EMVI	M	Surgery	nCRT
1	49	M	Caucasian	High rectum	Adenocarcinoma	T3	N0/2	T3	N1	+ve	No	Yes	No
2	71	F	Caucasian	High rectum	Adenocarcinoma	T3	N0	T3	N0	-ve	No	Yes	No
3	76	F	Caucasian	High rectum	Adenocarcinoma	T3	N1	T3	N1	+ve	No	Yes	No
4	79	F	Caucasian	Mid rectum	Tubulovillous adenoma with high grade dysplasia	T3	N0	N/R	N0	-ve	No	Yes	No
5	77	F	Caucasian	Low rectum	Tubulovillous adenoma with low grade dysplasia	T4	N0	N/R	N/R	-ve	No	Yes	No
6	74	F	Caucasian	High rectum	Adenocarcinoma	T2/3	N0	T2	N0	-ve	No	Yes	No
7	71	F	Caucasian	Low rectum	Adenocarcinoma	T2	N0	T2	N1	+ve	No	Yes	No

8	67	M	Caucasian	High rectum	Adenocarcinoma	T3	N2	N/R	N/R	NR	No	Yes	Yes
9	63	F	Caucasian	Low rectum	Dysplasia ("likely too superficial")	T4	N0	N/R	N/R	NR	No	Yes	Yes
10	66	F	Caucasian	High rectum	Adenocarcinoma	T3a	N1	T2	N0	NR	No	Yes	No
11	78	M	Caucasian	Low rectum	Adenocarcinoma	T2/3	N0	N/R	N/R	-ve	No	Yes	No
12	56	M	Caucasian	Low rectum	Adenocarcinoma	T3	N1	N/R	N/R	N/R	No	Yes	Yes
13	73	M	Caucasian	High rectum	Adenocarcinoma	N/R	N/R	T1	N1	-ve	No	Yes	No
14	77	M	Caucasian	Low rectum	Adenocarcinoma	T2/3	N1	T1	N0	-ve	No	Yes	No
15	45	M	Caucasian	High rectum	Adenocarcinoma	T3	N1	N/R	N/R	+ve	No	Yes	Yes

Table 3.3 Clinical and histopathological characteristics of all patients.

3.3.3 Global hypomethylation in tumour samples

A combined analysis of all 424,413 probes on the Illumina 450K array passing stringent QC metrics showed that global levels of DNA methylation are decreased in rectal tumour samples compared with normal unaffected samples (rectal tumour mean β = 0.468, normal mucosa β = 0.471, mean difference = 0.003 P= 0.006). Therefore, global hypomethylation in rectal tumour samples vs normal mucosal samples was demonstrated. For each CpG site a β-value was generated by calculating the ratio of the normalised signal from the methylated probe to the sum of the normalised signals of the methylated and unmethylated probes. Consequently, the β-value is a measurement ranging continuously from zero (unmethylated) to one (fully methylated). Plotting the density of all β-values for all samples showed that, as expected, the data followed a bimodal distribution (Figure 3.3). The methylated peak for the normal mucosal samples was skewed to the left of the rectal tumour samples, corresponding to the higher global methylation levels of the normal mucosal samples. Given the global DNA methylation changes observed we also tested whether there was a difference in derived "DNA methylation" age between tumour and normal samples. We did not identify a significant difference suggesting there is no elevated aging in the tumour tissue (p=<0.05).

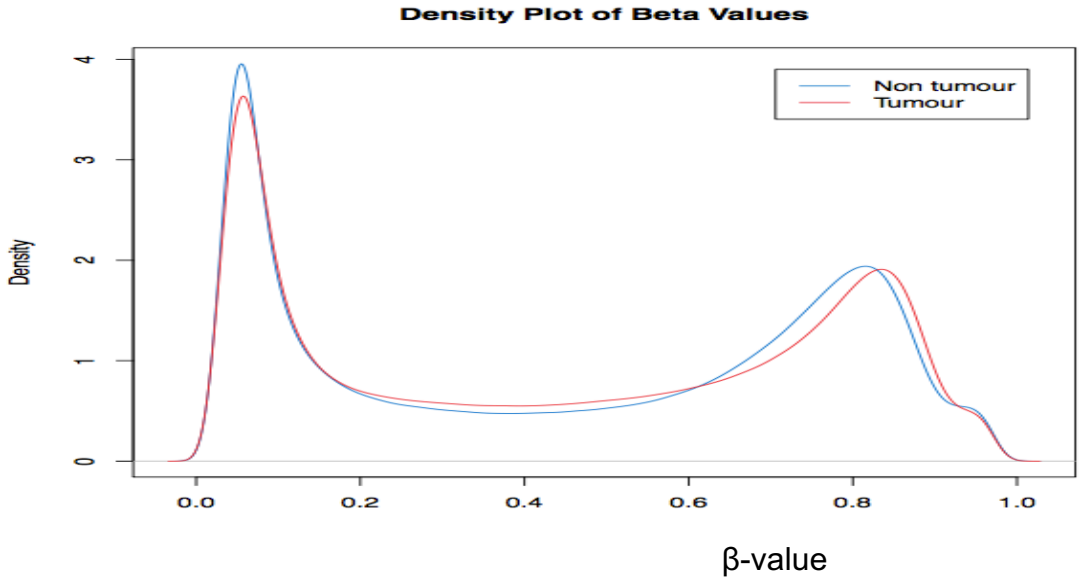


Figure 3.3 Beta density plot for rectal tumour and adjacent mucosal samples. A clear shift to the left is seen for normal mucosal samples. (Note the shift is on sites already methylated, as expected, with no difference on unmethylated sites).

3.3.4 Identification and distribution of differentially methylated positions (DMPs)

A linear regression model was used to identify DNA methylation at each probe associated with rectal cancer. Probes were identified as significantly associated if they passed a stringent threshold of P < 2E-7. We found that DNA methylation at autosomal sites was variable whereas overall there is a decreased global DNA hypomethylation in rectal cancer (Section 3.3.3).

Out of the 424,413 probes that passed the QC pipeline, a total of 176 probes were differentially methylated between rectal cancer tissues and normal mucosa. We refer to these as differentially methylated positions (DMPs).

The distribution of DMPs is relatively consistent across most autosomal chromosomes (Figure 3.4), however, one chromosome showed a notable enrichment of significant sites: 8.5% of CpG probes on chromosome 20 were identified as DMPs (enrichment = 4.12, p = 1.05E05) (Table 3.4).
Although the DMPs are seen throughout the genome, they are not equally distributed with regard to annotated genetic features. The hypermethylated and hypomethylated DMP sites were distributed over six gene categories: TSS1500 (15.3%), TSS200 (15.9%), 5′UTR (26.1%), 1st exon (19.9%), gene body (21%), 3′UTR (0.57%) and intergenic regions (22.7%) (Figure 3.5).
The distribution of the hypermethylated and hypomethylated DMPs over CpG islands (CGIs), CGI shores, CGI shelves and open sea regions was also analysed. As expected, the vast majority of CpG sites (79.0%) were located in the CpG islands, followed by CGI shores (10.8%), CGI Sea (9.1%) and CGI shelves (1.1%) (Figure 3.5).
For categorisation, the CpG counts were normalised by the number of CpGs in the same category represented on the 450K array, and the percentage of normalised CpG counts is indicated through the bars (Figure 3.6).
Compared to the distribution of all probes included in the dataset, DMPs were significantly enriched in promoter regulatory regions including CpG islands (percentage of significant probes = 79%, relative enrichment = 7.9, p = 3.63E-37), 5’ untranslated regions (5’ UTRs) (percentage of significant probes= 26.1%, relative enrichment = 2.14, p = 3.06E-05), genic first exons (percentage of significant probes = 19.9%, relative enrichment = 2.69, p = 1.89E-06).

Conversely, DMPs were significantly underrepresented in CGI shores (percentage of significant probes = 10.8%, relative enrichment = 0.39, p = 2.45E-05), CGI shelves (percentage of significant probes = 1.14%, relative enrichment = 0.11, p = 1.16E-05), CGI seas (percentage of significant probes = 9.1%, relative enrichment = 0.19, p = 2.36E-15) and gene bodies (percentage of significant probes = 21%, relative enrichment = 0.46, p = 9.95E-06) and 3' untranslated regions (percentage of significant probes = 0.57%, relative enrichment = 0.13, p = 0.008).

Relative enrichment and depletion in the vicinity of annotated transcription start site TSS200 and TSS1500 respectively was non-significant (percentage of probes in TSS200 = 16%, relative enrichment = 1.19, p = 0.38 and percentage of probes in TSS1500 = 15.3%, relative enrichment = 0.83, p = 0.43) (Figure 3.7). Table 3.5 and Table 3.6 show an overview of the distribution of all probes and DMPs across CpG and genomic features**.** Table 3.7 and Table 3.8 show the distribution of hypomethylated and hypermethylated probes, respectively, across CpG regions. Table 3.9 and Table 3.10 show the distribution of hypomethylated and hypermethylated probes, respectively, across genomic features.

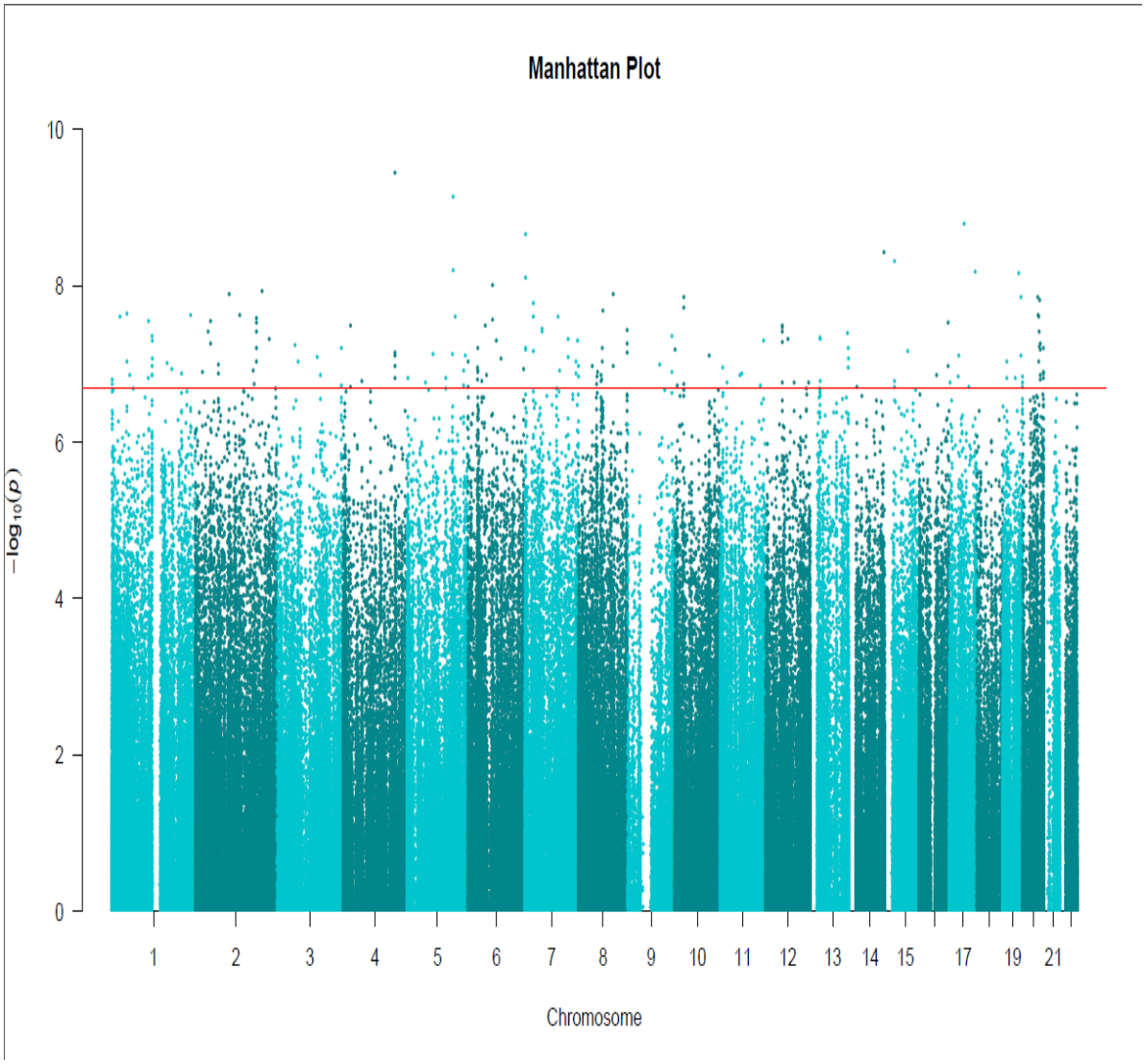


Figure 3.4 Manhattan plot showing the distribution of tumour associated DMPs across all autosomes. The red line represents the threshold (P < 1.25E-7). The dots above the red line portray the DMPs and the dots below the line depict all the probes on each chromosome.

	All	%	Rectal	%	P Value	Enrichment
Total	424413	100	176	100		
Chr1	41386	9.751350689	18	10.22727273	0.799228153	1.054406535
Chr2	30645	7.220561104	17	9.659090909	0.240852363	1.374025428
Chr3	22400	5.277877916	6	3.409090909	0.395350087	0.63331765
Chr4	17753	4.182953868	8	4.545454545	0.707256146	1.090815528
Chr5	21412	5.045085801	11	6.25	0.487208025	1.254879621
Chr6	31336	7.383374213	14	7.954545455	0.772132497	1.084051903
Chr7	25538	6.017252063	18	10.22727273	0.025491392	1.779909941
Chr8	18312	4.314665196	15	8.522727273	0.013491585	2.067048207
Chr9	8625	2.032218617	3	1.704545455	1	0.835907898
Chr10	21445	5.052861246	6	3.409090909	0.391046051	0.663102935
Chr11	25781	6.074507614	6	3.409090909	0.156041819	0.545605772
Chr12	21649	5.100927634	8	4.545454545	0.864719947	0.885874634
Chr13	10700	2.521129183	7	3.977272727	0.220431718	1.601871084
Chr14	13297	3.133033154	2	1.136363636	0.187490593	0.355282331
Chr15	13176	3.104523188	4	2.272727273	0.666420557	0.725754703
Chr16	18866	4.445198427	3	1.704545455	0.095782672	0.372666999
Chr17	24506	5.774092688	6	3.409090909	0.254086598	0.575839664
Chr18	5305	1.249961712	0	0	0.288840233	0
Chr19	22002	5.184101335	9	5.113636364	1	0.985669098
Chr20	9397	2.21411691	15	8.522727273	1.05E-05	4.119682271
Chr21	3731	0.87909654	0	0	0.41387867	0
Chr22	7286	1.71672404	0	0	0.080687631	0

Table 3.4 Distribution of DMPs split by autosomal chromosomes.

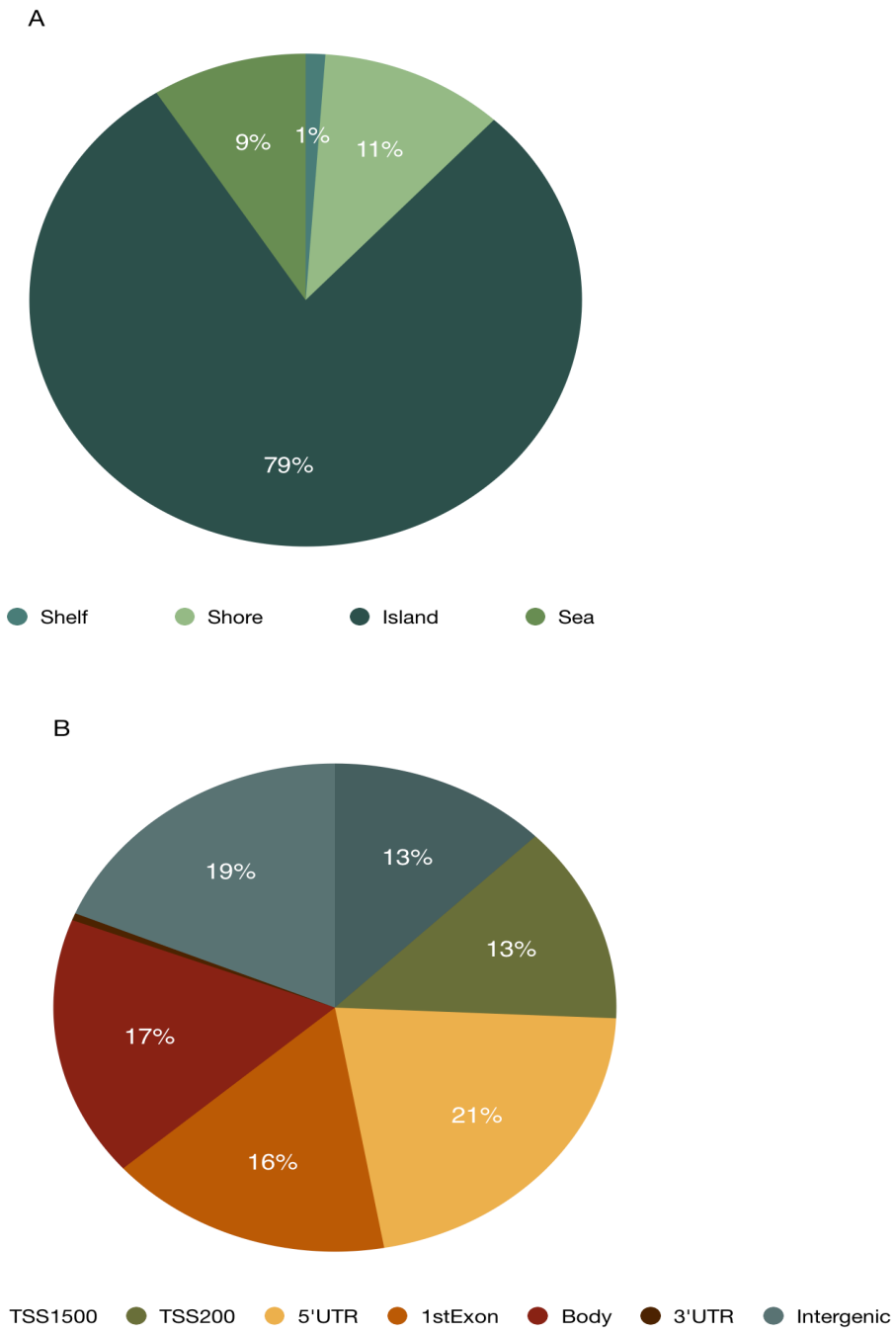


Figure 3.5 Pie charts showing the distribution of DMPs. A) across CpG sites B) across genomic features.

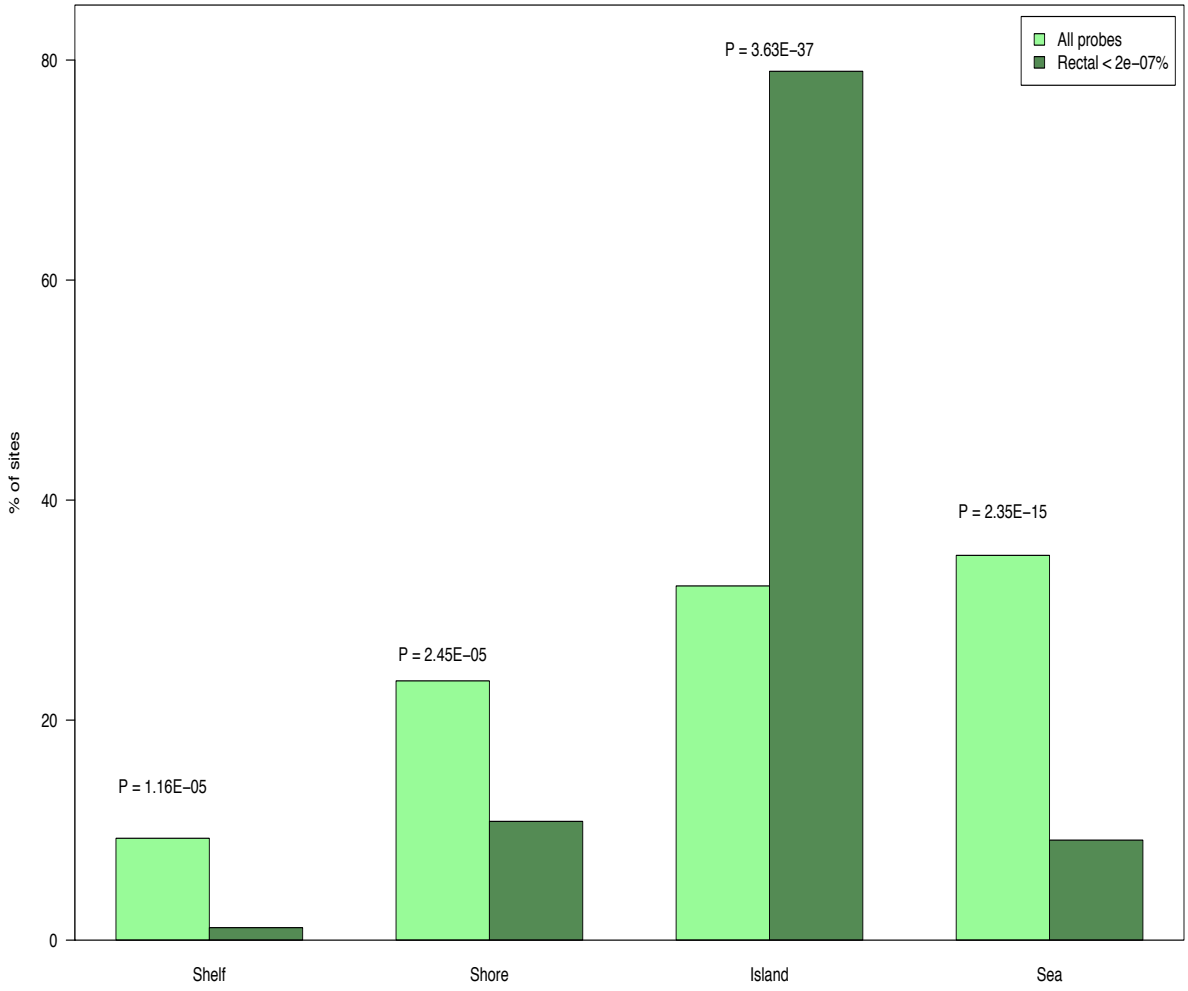


Figure 3.6 Bar chart depicting the distribution of probes across CpG sites. Light green bars indicate the distribution of all the probes whereas adjacent dark green bar indicates the percentage of the significant DMP for that same region. Above the bars is the p value of the difference.

	Total Probes	(%)	DMPs (p < 2E-07)	(%)	P Value	Enrichment
Total	424413	100	176	100		
Shelf	39297	9.26	2	1.14	1.16E-05	0.11
Shore	100025	23.57	19	10.80	2.45E-05	0.39
Island	136651	32.20	139	78.98	3.63E-37	7.92
Sea	148440	34.98	16	9.09	2.36E-15	0.19

Table 3.5 Overview of the distribution of the DMPs across CpG regions.

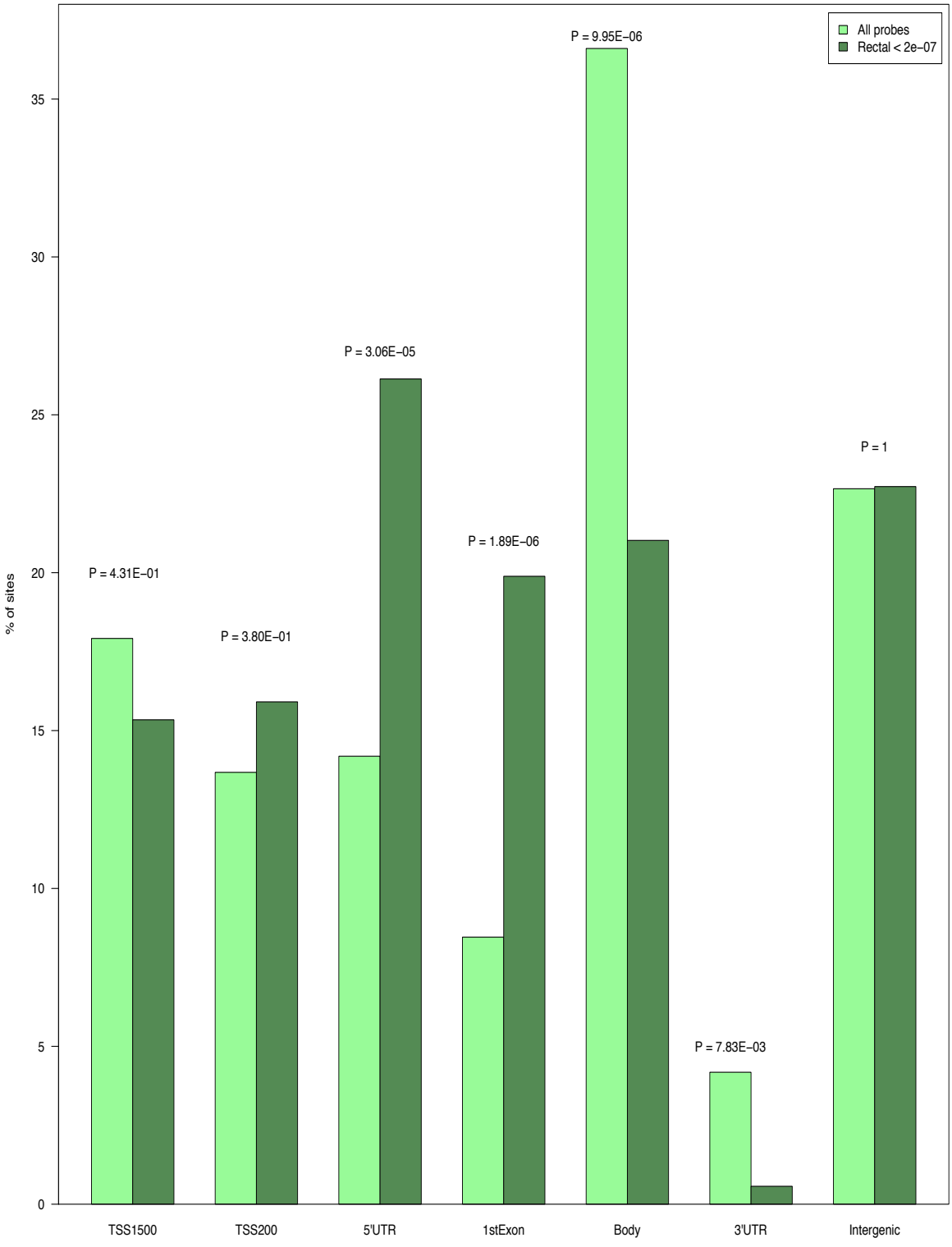


Figure 3.7. Bar chart depicting the distribution of probes across genomic features. Light green bars indicate the distribution of all the probes whereas adjacent dark green bar indicates the percentage of the DMP for that same region. Above the bars is the p value of the difference.

	All Probes	(%)	DMPs	(%)	P Value	Enrichment
Total	424413	100.00	176	100		
TSS1500	76045	17.92	27	15.34	0.431496981	0.83
TSS200	58046	13.68	28	15.91	0.380057727	1.19
5'UTR	60213	14.19	46	26.14	3.06E-05	2.14
1stExon	35899	8.46	35	19.89	1.89E-06	2.69
Body	155343	36.60	37	21.02	9.95E-06	0.46
3'UTR	17737	4.18	1	0.57	0.007827814	0.13
Intergenic	96166	22.66	40	22.73	1	1.00

Table 3.6 Overview of the distribution of the DMPS across genomic features.

	All Hypomethylated Probes	%	Significantly Hypomethylated Probes (p< 1.25E-7)	%	P Value	Enrichment
Total	197772	100	21	100		
Shelf	21769	11.00711931	1	4.761904762	0.72245339	0.404225178
Shore	41076	20.76937079	2	9.523809524	0.284586158	0.401521283
Island	48256	24.39981393	5	23.80952381	1	0.968230273
Sea	86671	43.82369597	13	61.9047619	0.122858355	2.083215505

Table 3.7 Overview of the distribution of hypomethylated probes across CpG regions.

	All Hypermethylated Probes	%	Significantly Hypermethylated Probes (p< 1.25E-7)	%	P value	Enrichment
Total	226641	100	155	100		
Shelf	17528	7.733816918	1	0.64516129	0.000112052	0.077394769
Shore	58949	26.009857	17	10.96774194	5.48E-06	0.35022811
Island	88395	39.00221054	134	86.4516129	2.47E-34	9.992248395
Sea	61769	27.25411554	3	1.935483871	2.69E-17	0.052635237

Table 3.8 Overview of the distribution of hypermethylated probes across CpG regions.

	All Hypomethylated Probes	%	Significantly Hypomethylated Probes (p< 1.25E-7)	%	P Value	Enrichment
Total	197772	100	21	100		
TSS1500	30491	15.42	3	14.29	1	0.91434221
TSS200	23335	11.80	0	0.00	0.164494492	0
5'UTR	24362	12.32	2	9.52	1	0.749249321
1stExon	14137	7.15	1	4.76	1	0.649439526
Body	73231	37.03	5	23.81	0.261909754	0.531426339
3'UTR	7947	4.02	0	0.00	1	0
Intergenic	54056	27.33	11	52.38	0.014264388	2.924888549

Table 3.9 Overview of the distribution of hypomethylated probes across genomic features.

	All Hypermethylated Probes	%	Significant Hypermethylated probes (p< 1.25E-7)	%	P Value	Enrichment
Total	226641	100	155	100		
TSS1500	45554	20.10	24	15.48	0.161287284	0.728144993
TSS200	34711	15.32	28	18.06	0.371135895	1.219233509
5'UTR	35851	15.82	44	28.39	9.14E-05	2.110903761
1stExon	21762	9.60	34	21.94	4.03E-06	2.64796571
Body	82112	36.23	32	20.65	3.59E-05	0.457714581
3'UTR	9790	4.32	1	0.65	0.016079783	0.143721678
Intergenic	42110	18.58	29	18.71	0.918071224	1.008588913

Table 3.10 Overview of the distribution of hypermethylated probes across genomic features.

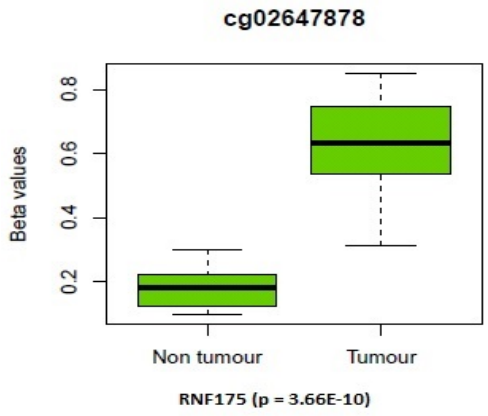
3.3.5 Identification of DMPs associated with rectal cancer

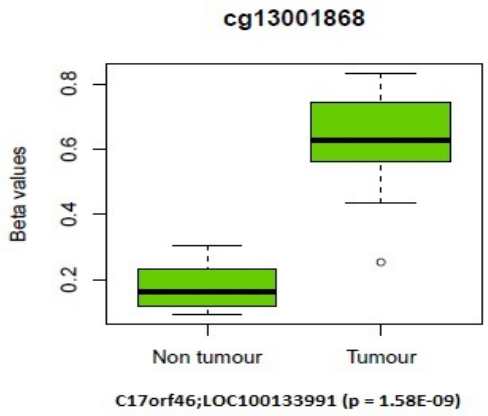
The top 10 autosomal DMPs (ranked by P value) are listed in Table 3.11. They consist of nine hypermethylated probes (characterised by higher methylation in tumour samples) and one hypomethylated probe (cg18538668, characterised by lower DNA methylation in tumour samples). Figure 3.8 shows box plots demonstrating the beta value difference of the top ten probes with gene name and p values. Differences in DNA methylation were observed across all of the rectal samples (n=30) between matched normal mucosa and tumour for the top ranked DMPs (Figure 3.9). A complete list of all 176 DMPs is found in Supplementary Table 2 (Appendix 4).

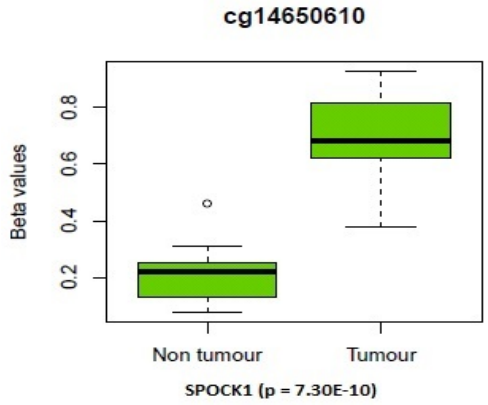
Out of the 176 DMPs, 21 were hypomethylated in tumour samples and the rest, 155 probes, were hypermethylated. Overall there is a highly significant enrichment of hypermethylated autosomal DMPs compared with hypomethylated DMPs (hypermethylated DMPs n = 155 P <2.2e-16). Figure 3.10 shows a heatmap of the DMPs identified.
Some DMPs have not been annotated to genes, leaving a total of 137 DMPs annotated to known genes (95 genes in total, 71%).

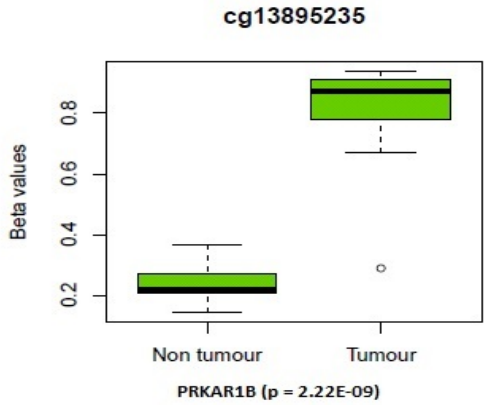
Probe	P Value	Regression coefficient	Chr	Position	Gene name	Probe location
cg02647878	3.66E-10	0.45	4	154681197	RNF175	5'UTR; 1stExon
cg14650610	7.30E-10	0.48	5	136834492	SPOCK1	5'UTR
cg13001868	1.58E-09	0.45	17	43339223	C17orf46; LOC100133991	Body; TSS1500
cg13895235	2.22E-09	0.57	7	752292	PRKAR1B	5'UTR
cg18538668	3.79E-09	-0.20	14	103839038		
cg03061682	4.83E-09	0.42	15	28352098		
cg24847829	6.29E-09	0.40	5	136834464	SPOCK1	5'UTR
cg26034516	6.48E-09	0.35	17	76228121	LOC283999	Body
cg03576469	6.83E-09	0.29	19	46917061	CCDC8	TSS200
cg18601167	8.02E-09	0.56	7	752286	PRKAR1B	5'UTR

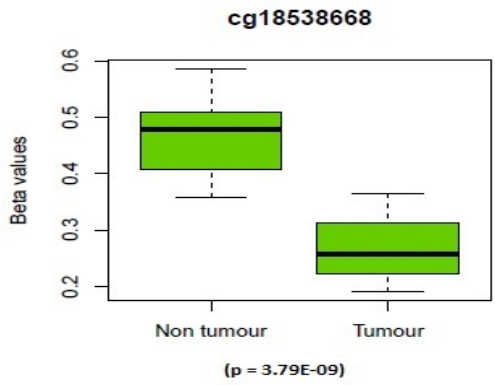
Table 3.11 The top ten DMPs. Table lists each probe, its chromosomal position and genic location, its annotated gene, p value and direction of effect

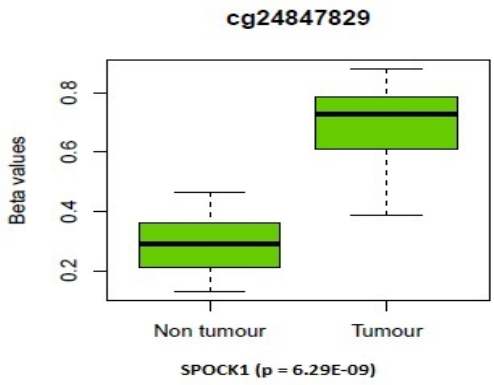

cg02647878
Beta values
0.2
0.4
0.6
0.8
Non tumour
Tumour
RNF175 (p = 3.66E-10)

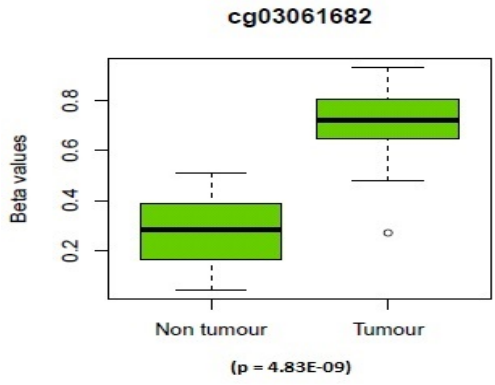

cg13001868
Beta values
0.2
0.4
0.6
0.8
Non tumour
Tumour
C17orf46;LOC100133991 (p = 1.58E-09)

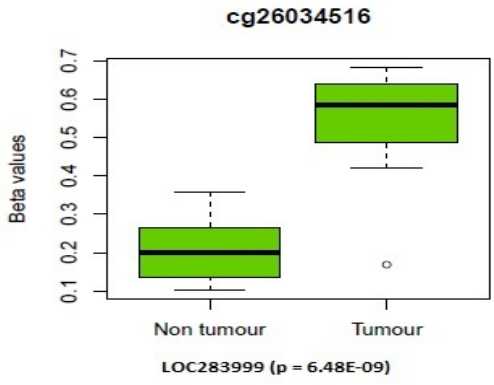

cg14650610
Beta values
0.2
0.4
0.6
0.8
Non tumour
Tumour
SPOCK1 (p = 7.30E-10)


cg13895235
Beta values
0.2
0.4
0.6
0.8
Non tumour
Tumour
PRKAR1B (p = 2.22E-09)


cg18538668
Beta values
0.2
0.3
0.4
0.5
0.6
Non tumour
Tumour
(p = 3.79E-09)


cg24847829
Beta values
0.2
0.4
0.6
0.8
Non tumour
Tumour
SPOCK1 (p = 6.29E-09)


cg03061682
Beta values
0.2
0.4
0.6
0.8
Non tumour
Tumour
(p = 4.83E-09)


cg26034516
Beta values
0.1
0.2
0.3
0.4
0.5
0.6
0.7
Non tumour
Tumour
LOC283999 (p = 6.48E-09)

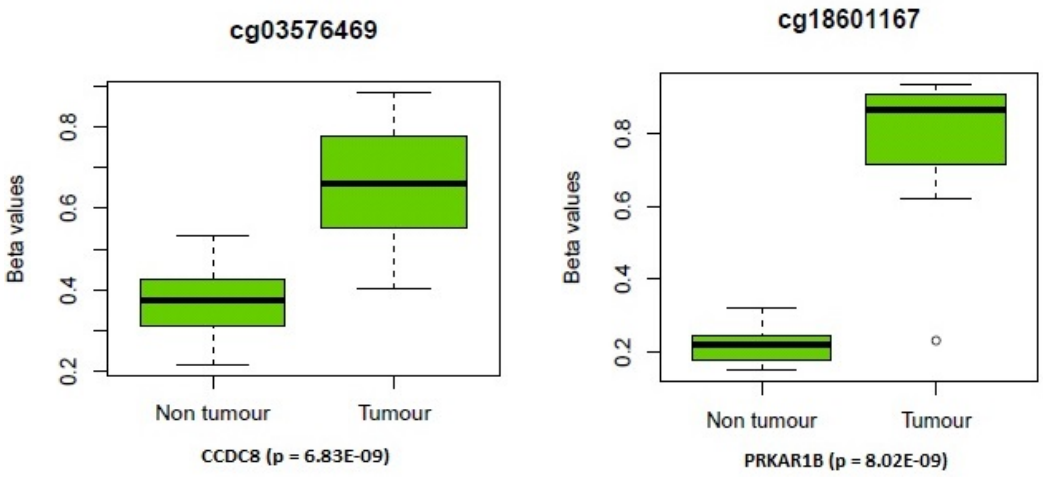
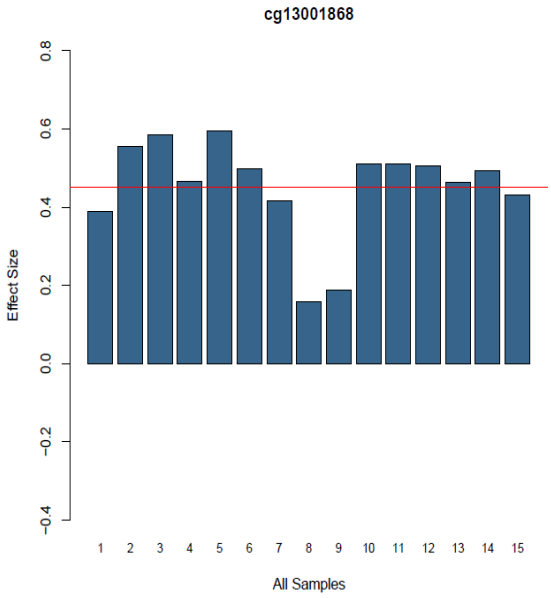
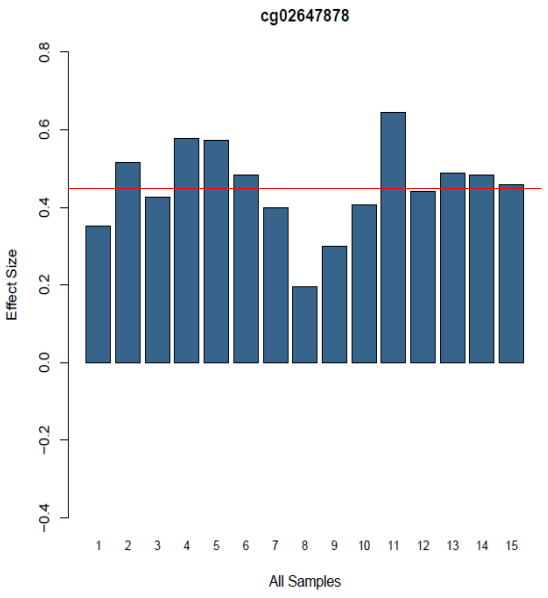
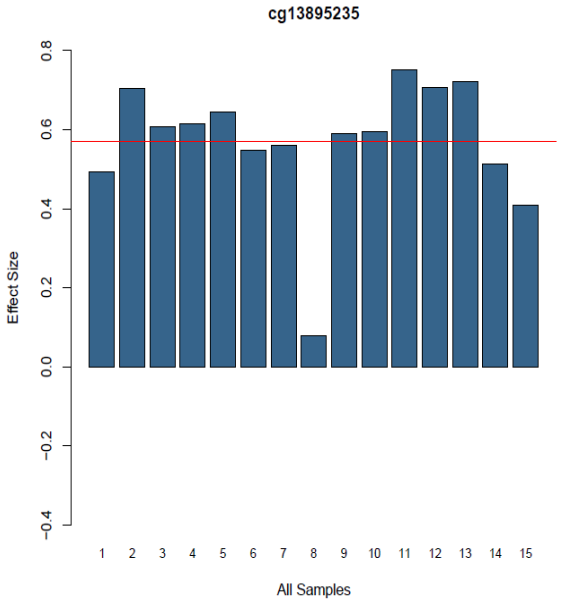
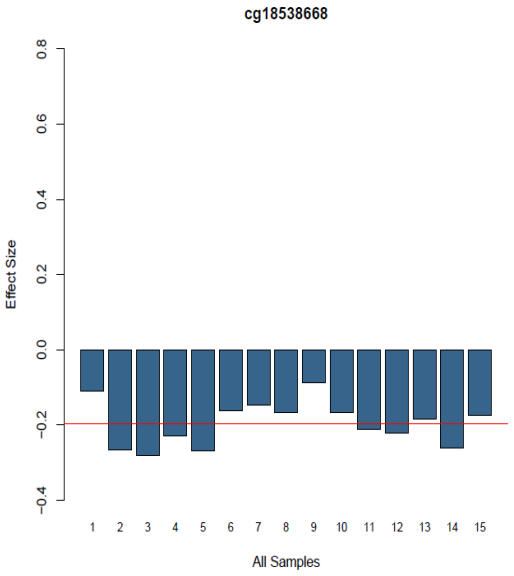


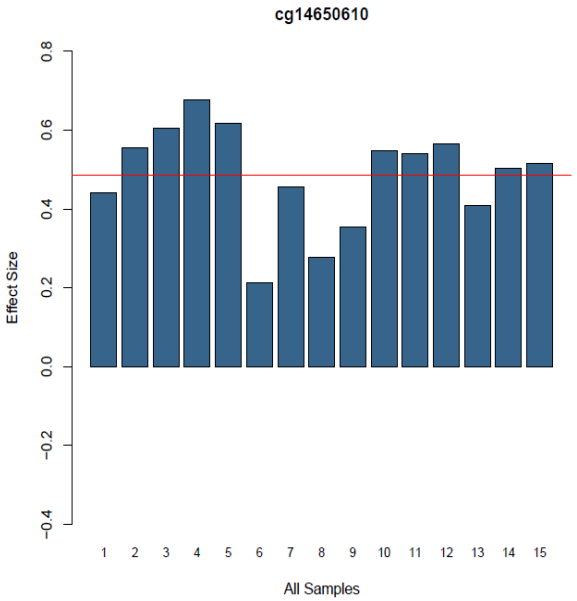
Figure 3.8 Box plots of beta values of the top ten probes. The box plots are showing the beta values of tumour vs non-tumour samples annotated with the probe name (at the top of plot) and gene name (if available) with p values in brackets (at the bottom of plot). One DMP (probe cg18538668) is more hypomethylated in tumour vs normal mucosa. All nine other DMPs are hypermethylated in tumour vs normal mucosa. Whiskers depict the 25 and 75% quartile of the data.

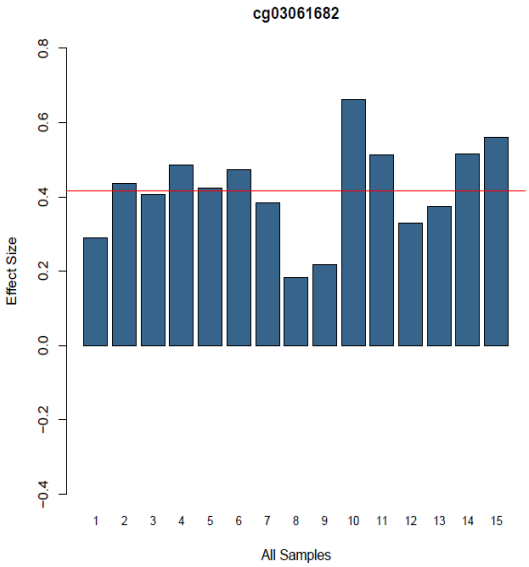

cg13001868
Effect Size
All Samples


cg02647878
Effect Size
All Samples


cg13895235
Effect Size
All Samples


cg18538668
Effect Size
All Samples


cg14650610
Effect Size
All Samples


cg03061682
Effect Size
All Samples

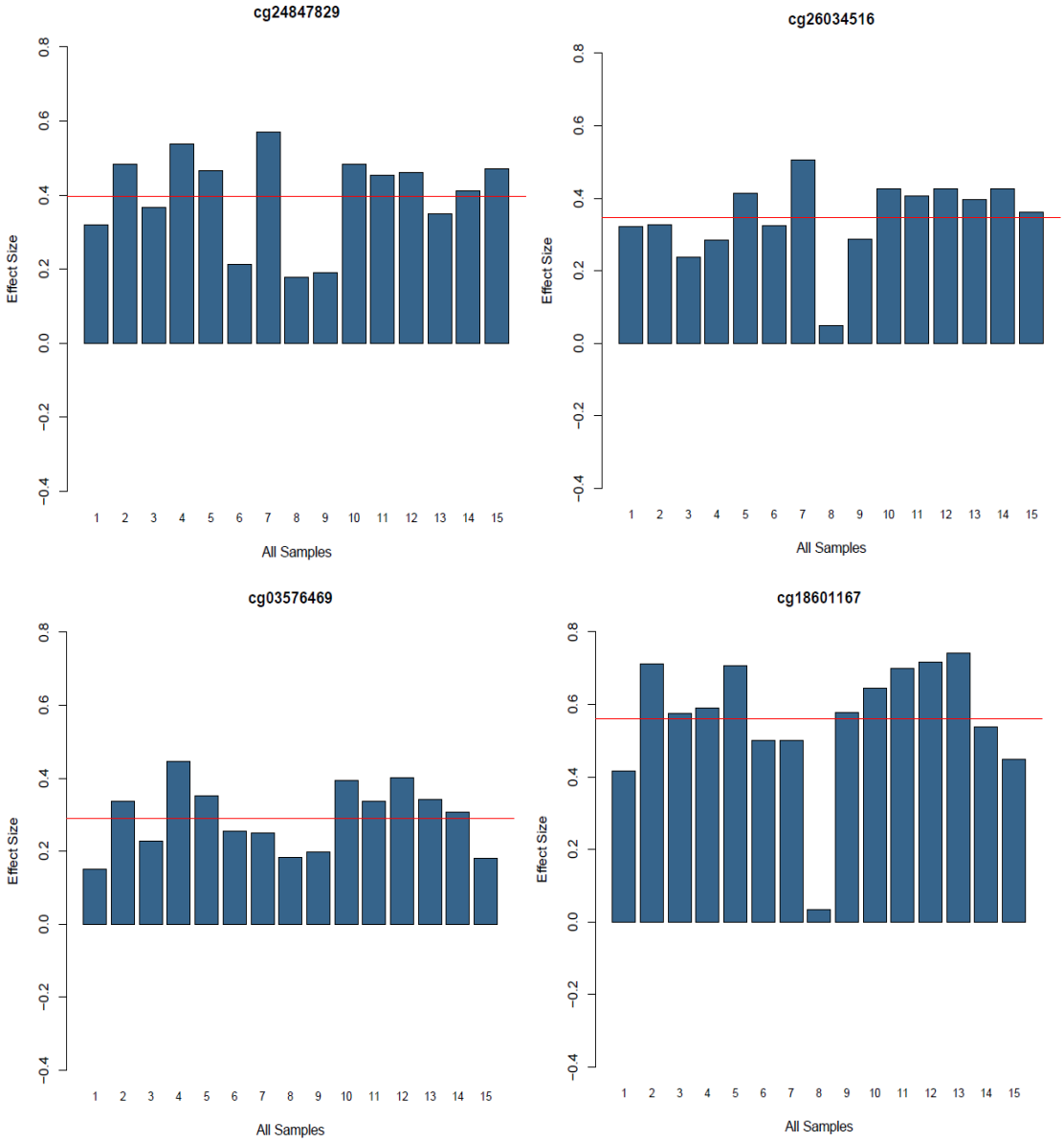


Figure 3.9 Top ten DMPs between each pair of rectal samples. Graphs showing the difference in DNA methylation (Δβ) between each pair of matched rectal samples (tumour and normal mucosa) for each of the ten top-ranked differentially methylated positions (DMPs). Mean Δβ across all 15 pairs is highlighted with red bar. Consistent within each pair are differences in DNA methylation. Notably, sample 8 consistently showed lower (Δβ) than the other samples across all DMPs.

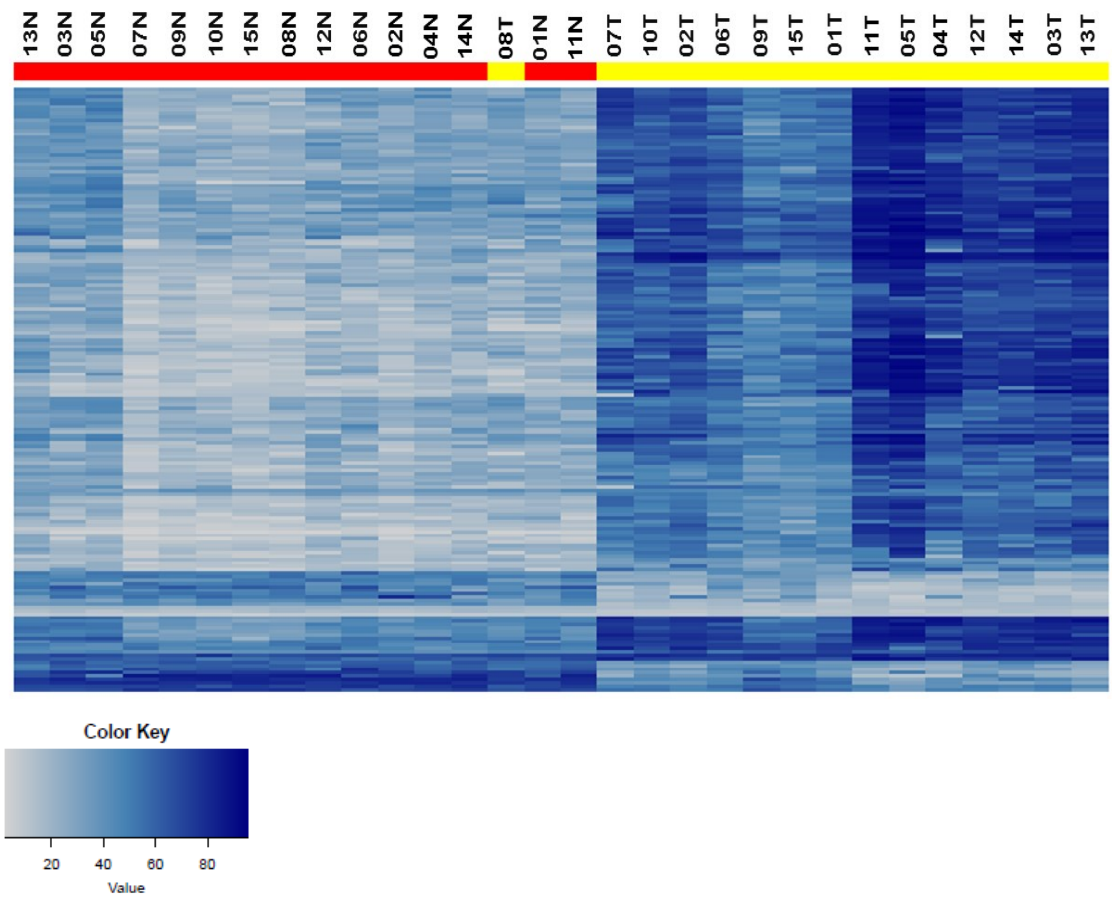


Figure 3.10 Heatmap of DMPs. This heatmap is depicting the difference in the top DMPs between normal rectal samples (red) and rectal tumour (yellow). The sample identifications are written at the top, with N for 'normal' and T for 'tumour'.

Several of our top DMPs are located in the vicinity of genes that have previously been implicated in cancer pathways. Of note:

cg02647878, the top ranked DMP, was hypermethylated in rectal tumour compared with normal mucosa (p =3.66E-10). It is located on chromosome 4, 189 kb downstream of the transcriptional start site (TSS) of the *RNF175* gene, which encodes ring finger protein 175 and is involved in the prevention of oncogenesis (Ng et al., 2003). Also, amongst the top ten ranked DMPs is the hypermethylated cg03576469 probe (p= 6.83E-09) associated with the gene CCDC8, a regulator of microtubule dynamics and associated with metastasis of breast cancer (Pangeni et al., 2015). Interestingly, two of the top ten ranked

DMPs were in the vicinity of PRKAR1B gene. The probes cg13895235 and cg18601167 are located on chromosome 7 downstream of the TSS of PRKAR1B gene (p = 2.22E-09 and p = 8.02E-09 respectively). Similarly, SPOCK1 gene was associated with two top ten-ranked DMPs: cg14650610 and cg24847829 (p = 7.30E-10 and p = 6.29E-09 respectively), located on chromosome 5 downstream of the TSS of SPOCK1. Both SPOCK1 and PRKAR1B play a role in cancer pathways (Miao et al., 2013) (Naviglio et al., 2009). Table 3.12 lists all of the DMPs (p < 2E-7) identified in the analysis that have previously been found to play a role in cancer, along with the genes they are annotated to

Probe(s)	Genomic Position	Illumina Gene Annotation	Genic Region	DNA Methylation Difference (%)	P Value	Gene Function	Literature Finding
cg20295442	Chr8:67344665	ADHFE1	TSS200	47.80	1.37E-07	Encodes hydroxyacid-oxoacid transhydrogenase	**Frequently methylated in CRC. Methylation-mediated down expression of *ADHFE1* and proliferation of CRC cells alcohol induced (Moon et al., 2014)**
cg20912169	Chr8:67344720	ADHFE1; ADHFE1	5'UTR; 1stExon	47.99	1.64E-07		
cg26328335	Chr12:50354840	AQP5	TSS1500	44.60	3.26E-08	Encodes Aquaporin 5, a water channel protein.	**Associated with CRC metastasis (Kang et al., 2008) and with tumour TNM stage, lymph node and distant metastases (Shan et al., 2014)**
cg08266366	Chr12:50354998	AQP5	TSS1500	37.51	3.38E-08		
cg03020208	Chr12:50354962	AQP5	TSS1500	30.40	3.85E-08		
cg15336765	Chr12:50355307	AQP5; AQP5	1stExon; 5'UTR	42.95	5.15E-08		
cg13850380	Chr1:1475143	C1orf70	Body	31.76	1.60E-07	Encodes Chromosome 1 open reading frame 70	**Methylated in CRC (Naumov et al., 2013)**
cg16306898	Chr1:1475675	C1orf70	1stExon	47.78	1.83E-07		

cg17170568	Chr7:156433406	C7orf13;	TSS200;	17.71	5.05E-08	Encodes Chromosome 7 open reading frame 13	**Methylation levels implicated in expression levels and clinical outcomes in glioblastoma (Etcheverry et al., 2010)**
cg13405887	Chr9:132382812	C9orf50	1stExon	42.46	4.37E-08	Encodes Chromosome 9 open reading frame 50	**Novel blood based DNA methylation biomarker for CRC early detection (Naumov et al., 2013)**
cg14015706	Chr9:132382433	C9orf50	1stExon	50.26	1.30E-07		
cg03576469	Chr19:46917061	CCDC8	TSS200	29.07	6.83E-09	A regulator of microtubule dynamics	**Associated with metastasis of breast cancer (Pangeni et al., 2015)**
cg08516516	Chr5:115152492	CDO1	TSS200	30.09	1.50E-07	Encodes Cysteine dioxygenase type 1 which is involved in metabolic pathways of hypotaurine and pyruvate	**Methylation plays a role in CRC, is a promising biomarker for CRC detection (Yamashita et al., 2014)**
cg11573679	Chr2:68546467	CNRIP1	1stExon	47.54	1.03E-07	Encodes a protein that interacts with the C-terminal tail of cannabinoid receptor 1.	**Methylated CNR1P1 was recently included in a novel combined biomarker panel for detection of CRC and**

cg24171907	Chr2:68546579	CNRIP1; CNRIP1	1stExon; 5'UTR	44.70	1.32E-07		**premalignant melanomas (Lind et al., 2011)**
cg21938148	Chr13:110958977	COL4A1; COL4A1	Body; TSS1500	47.95	4.01E-08	Encodes a type IV collagen alpha protein.	**Methylated in peripheral blood DNA in CRC cancer cases (Mitchell et al., 2014)**
cg11751707	Chr2:38302587	CYP1B1	5'UTR	27.31	3.85E-08	Encodes cytochrome P450 1B1	**Methylation levels involved in development of some CRCs and altered response to chemotherapy in CRC (Habano et al., 2009)**
cg17393267	Chr3:192127356	FGF12; FGF12	TSS1500; Body	28.74	6.15E-08	Encodes Fibroblast growth factor 12	**Hypermethylated in CRC, promising early detection biomarker (Li et al., 2012)**
cg16366473	Chr3:192126849	FGF12; FGF12	TSS200; Body	41.24	1.88E-07		
cg26958524	Chr16:86613067	FOXL1	1stExon	33.70	2.90E-08	Encodes Forkhead box protein L1	**Hypermethylated in adenoid cystic carcinoma of salivary gland (Bell et al., 2012)**
cg03424342	Chr3:120169783	FSTL1; FSTL1	5'UTR; 1stExon	27.91	8.35E-08	Encodes Follistatin-related protein 1	**Reduction in DNA methylation after treatment with Celecoxib in oesophageal carcinoma (Liu JF, 2016)**

cg24820783	Chr10:26504969	GAD2	TSS1500	24.71	1.41E-08	Encodes Glutamate Decarboxylase 2	**Hypermethylated in CRC (Li et al., 2012)**
cg11903130	Chr10:26506751	GAD2	Body	34.81	1.82E-07		
cg11328303	Chr10:26505440	GAD2; GAD2	5'UTR; 1stExon;	27.11	1.94E-08		
cg13389502	Chr17:1961440	HIC1	Body	16.17	1.72E-07	Growth regulatory and tumour suppressor gene	**Hypermethylated in CRC, the loss of a single HIC1 allele was found to accelerate polyp formation in mice with APC gene mutation (Mohammad et al., 2011) Hypermethylation was also found to play an important role in the initiation of the disease (Wahab et al., 2011)**
cg18607529	Chr7:50343869	IKZF1	TSS1500	40.88	3.47E-08	Encodes IKAROS Family Zinc Finger 1 protein	**Methylated in several cancers, potential to be used as a blood biomarker in the detection of CRC recurrence (Pedersen et al., 2015)**
cg07589773	Chr7:50343883	IKZF1	TSS1500	37.26	3.80E-08		
cg26415547	Chr12:66583048	IRAK3; IRAK3	1stExon; 5'UTR	31.99	4.82E-08	Encodes interleukin-1 receptor-	**Methylated in hepatocellular carcinoma tissues, associated with**

						associated kinase 3 protein	**tumour stage and poor prognosis (Kuo et al., 2015)**
cg17228900	Chr6:391764	IRF4; IRF4	5'UTR; 1stExon	45.37	9.42E-08	Belongs to the Interferon Regulatory factor family of transcription factors which contribute to the regulation of the interferon signalling pathway IFN	**Plays a role in rectal cancer (Slattery et al., 2011)**
cg25024074	Chr2:182322501	ITGA4	1stExon	31.15	2.94E-08	Encodes Integrin Alpha 4 receptor for fibronectin	**Frequently methylated in CRC and adenoma tissues (Zhang et al., 2015)**
cg11947981	Chr2:182322749	ITGA4	Body	39.48	6.39E-08		
cg06952671	Chr2:182322268	ITGA4; ITGA4	5'UTR; 1stExon	47.21	2.55E-08		
cg20415809	Chr2:182321855	ITGA4; ITGA4	5'UTR; 1stExon	38.61	3.92E-08		
cg21995919	Chr2:182322279	ITGA4; ITGA4	5'UTR; 1stExon	41.46	9.27E-08		
cg05447008	Chr6:73331114	KCNQ5	TSS1500	40.04	9.87E-09	Encodes Potassium Voltage-Gated Channel Subfamily Q Member 5	**Hypermethylated in CRC (Ashktorab et al., 2013)**
cg04377145	Chr6:73331191	KCNQ5	TSS1500	32.90	2.72E-08		

cg23977631	Chr2:100938799	LONRF2; LONRF2	5'UTR; 1stExon	45.66	1.30E-08	Encodes LON Peptidase N-Terminal Domain and Ring Finger 2	**Hypermethylation thought to contribute to tumourigenesis or rectal cancer (Hua et al., 2017)**
cg27200446	Chr6:41606439	MDFI	5'UTR	57.21	1.62E-07	Encodes MyoD Family Inhibitor	**Hypermethylated in CRC (Lin et al., 2014)**
cg16964348	Chr7:24323799	NPY	TSS200	39.09	1.65E-08	Encodes Neuropeptide Y protein	**Hypermethylated in CRC and promising biomarker (Roperch et al., 2013)**
cg24242823	Chr7:24323675	NPY	TSS200	37.21	6.79E-08		
cg25884711	Chr7:24323840	NPY	5'UTR; 1stExon	41.79	2.52E-08		
cg04366687	Chr8:145107199	OPLAH	Body	35.31	7.26E-08	Encodes 5-oxoprolinase, an enzyme responsible for glutathione synthesis and degradation	**Hypermethylated in CRC (Naumov et al., 2013)**
cg13895235	Chr7:752292	PRKAR1B	5'UTR; TSS200	56.92	2.22E-09	A regulatory subunit of cyclic AMP dependent protein kinase A (PKA)	**Involved in the cAMP/PKA signalling pathway which is altered in different cancers and may be used in cancer therapy and diagnosis (Naviglio et al., 2009)**
cg18601167	Chr7:752286	PRKAR1B	5'UTR; TSS200	55.98	8.02E-09		
cg12441126	Chr7:751962	PRKAR1B	5'UTR	41.41	6.21E-08		
cg20381963	Chr7:752238	PRKAR1B	5'UTR; TSS200	50.90	6.48E-08		

cg24645214	Chr8:54789978	RGS20	Body	30.08	1.45E-07	Regulator of G protein signalling	**Upregulated in a variety of tumours, may promote metastasis of tumour cells (Yang et al., 2016)**
cg02647878	Chr4:154681197	RNF175; RNF175	5'UTR; 1stExon	44.94	3.66E-10	Encodes ring finger protein 175	**Involved in the ubiquitination pathway, thought to contribute to the aetiology of tumours (Hoeller et al., 2006)**
cg18355902	Chr4:154681128	RNF175; RNF175	5'UTR; 1stExon	40.11	7.28E-08		
cg01440841	Chr4:154681066	RNF175; RNF175	5'UTR; 1stExon	35.12	1.06E-07		
cg25645268	Chr4:154710598	SFRP2	TSS1500	30.58	7.81E-08	Encodes secreted frizzled-related protein 2	**Frequently methylated in CRC and adenoma tissues (Zhang et al., 2015)**
cg03064067	Chr12:85306916	SLC6A15	TSS1500	39.48	1.70E-07	A disintegrin-like and metallopeptidase (reprolysin type) with thrombospondin type 1 motif)	**Hypermethylated in CRC (Kim et al., 2011)**
cg14658804	Chr5:168728213	SLIT3	TSS200	33.08	1.92E-07	Encodes Slit Guidance Ligand 3	**Frequently methylated in CRC and significantly associated with TNM stage, lymph node metastasis and differentiation (Huang et al., 2015)**
cg09073398	Chr5:168727762	SLIT3; SLIT3	5'UTR; 1stExon	38.52	1.20E-07		

cg24190603	Chr6:84418433	SNAP91; SNAP91	Body; 5'UTR	35.53	5.05E-08	Encodes Synaptosome Associated Protein 91	**Downregulated due to promoter hypermethylation in around 60% of colorectal cancer and may be associated with treatment resistance (Beggs et al., 2013)**
cg09296001	Chr7:127672564	SND1	Body	52.17	8.21E-08	Staphylococcal nuclease homology domain 1	**Frequently up-regulated in human colon cancers (Naumov et al., 2013)**
cg12628196	Chr7:127672458	SND1; LRRC4	Body; TSS1500	40.31	4.87E-08		
cg14650610	Chr5:136834492	SPOCK1	5'UTR	48.50	7.30E-10	Encodes the protein core of a seminal plasma proteoglycan containing chondroitin- and heparan-sulfate chains.	**Promotes tumour growth and metastasis in prostate cancer (Chen et al., 2016), contributes to metastases of recurrent glioblastoma (Yu et al., 2016) and plays a role in gallbladder and lung cancers (Miao et al., 2013)**
cg24847829	Chr5:136834464	SPOCK1	5'UTR	39.69	6.29E-09		
cg12619536	Chr1:108508067	VAV3	TSS1500	41.58	1.49E-07	A member of the VAV gene family	**Overexpressed in CRC and is correlated with tumour metastases and invasion (Uen et al., 2015)**

cg25480336	Chr20:50720908	ZFP64	Body	38.50	1.50E-08	Encodes Zinc Finger Protein 64	**Methylation levels indicate different molecular mechanisms for tumourigenesis in laterally spreading tumours of the colorectum (Sugai et al., 2016)**
cg17892556	Chr19:12267464	ZNF625; ZNF625	1stExon; 5'UTR	46.62	1.49E-07	Encodes Zinc Finger Protein 625	**Hypermethylated in CRC (Lin et al., 2014)**

Table 3.12 Summary of DMPs involved in rectal cancer, CRC, cancer pathways or colorectal diseases.

3.3.6 Identification of differentially methylated regions (DMRs) associated with rectal cancer

Many DMPs do not represent isolated changes at specific sites but rather occur in clusters known as differentially methylated regions (DMRs). Region level analysis was performed using *Comb-p* (Section 2.6.2.3) to identify spatially correlated regions of differential DNA significantly associated with rectal cancer. We allowed a maximum distance 500 bp between neighbouring CpG sites, and only included probes with a P-value <1e-7 in the initial epigenome wide association scan as starting points for identifying potential DMRs. For each DMR, we report the combined P, which is Stouffer–Liptak–Kechris-corrected for regional correlation structure, and the multiple-testing-corrected Šidák P-value. The latter corrects the combined P for na/nr tests, where na is the total number of probes tested in the initial epigenome wide association scan and nr the number of probes in the given region.

In total, 828 differentially DMRs were identified in rectal cancer vs normal tissue (P < 1E-7, number of probes >=5). The top ten ranked DMRs (ranked by Sidak-corrected P value) are shown in Table 3.13 and the top region is shown in Figure 3.11. 53 of the genes annotated to DMRs were also annotated to DMPs. They are listed in Supplementary Table 3 (Appendix 4). A complete list of all significant DMRs is found in Supplementary Table 4 (Appendix 4).

Six of the genes annotated to our top ten DMRs (Table 3.13) were found to be of interest, of note:

- The top ranked DMR was annotated to the gene PRKAR1B (p<1.24E-19) which was also listed twice in the top ten DMPs (Table 3.4), is a gene involved in signalling pathway which are often altered in different cancers and may be used in cancer therapy and diagnosis (Naviglio et al., 2009).
- ITGA4 gene, also annotated to several DMPs, is frequently hypermethylated in CRC (Table 3.5).
- Hypermethylated TBX15, has been correlated with low expression of the TBX15 protein and thought to play a role in the pathogenesis of ovarian cancer (Gozzi et al., 2016).

- GSTM2 methylation has previously been identified in prostate cancer, breast cancer and oral squamous cell carcinoma (Ashour et al., 2014, Rauscher et al., 2015, Li YF et al., 2015) and considered a potential biomarker for rectal cancer (Wei et al., 2016)
- NPY methylation has been successfully included in a six-CpG methylation biomarker panel for prediction of treatment (CRT) response in patients with oesophageal squamous cell carcinoma (Chang et al., 2016).
- EYA4 promoter methylation has previously been linked to CRC and shown to function as a tumour suppressor gene as well as being a potential biomarker (Kim SJ et al., 2015)

Furthermore, six genes - CDH13, RUNX3, SFRP1, SFRP2, ESR1 and ITGA4 - known to play a role in the in the multi-step process from normal colonic epithelium leading to adenocarcinoma (Section 1.6.3.3), were annotated to DMRs in our rectal cancer samples. Table 3.14 lists their function and Figure 3.12 illustrates their role in the progression of normal colon to adenocarcinoma.

Other DMRs identified were also in the vicinity of genes that have previously been implicated in colon and rectal cancer. The SEPT9 gene for example, a member of the Septin family of genes, is a well-studied and promising biomarker for CRC (Grutzman et al., 2008, Tanzer et al., 2010). Genes SFRP1, SFRP2 and Fli1 are also considered potential biomarkers for detection and monitoring of colon and rectal cancer (Zhang et al., 2007, Vymetalkova et al., 2016, Nagaska at al., 2009, Tang et al., 2011). Table 3.15 lists all of the genes annotated to DMRs that have been implicated in rectal cancer or considered to be a potential biomarker for cancer treatment response.

Chromosome	Start	End	Gene	N of Probes	Sidak P Value	CpG region
chr7	751830	752456	PRKAR1B	9	1.24E-19	island
chr3	129693370	129694683	TRH	11	7.14E-19	island
chr1	119526783	119528848	TBX15	14	8.02E-19	island
chr10	118031632	118033902	GFRA1	24	8.33E-19	island
chr2	182321489	182322841	ITGA4	11	3.88E-18	island
chr7	24323128	24323939	NPY	10	1.82E-17	island
chr8	69243285	69244734	C8orf34-AS1; C8orf34	12	1.66E-16	island
chr6	133561224	133562776	EYA4	41	1.76E-16	island
chr5	134362967	134363973	PITX1	9	2.38E-16	island
chr8	70982867	70983760	PRDM14	10	3.32E-16	island

Table 3.13. The top ten DMRs. Chromosomal location is indicated in the first three columns, followed by gene name, the number of probes identified, the corrected Sidak p value and the CpG region in the last column. As expected, all methylated genes were located in the CpG islands. The table with all the DMRs is found in Supplementary Table 4 (Appendix 4).

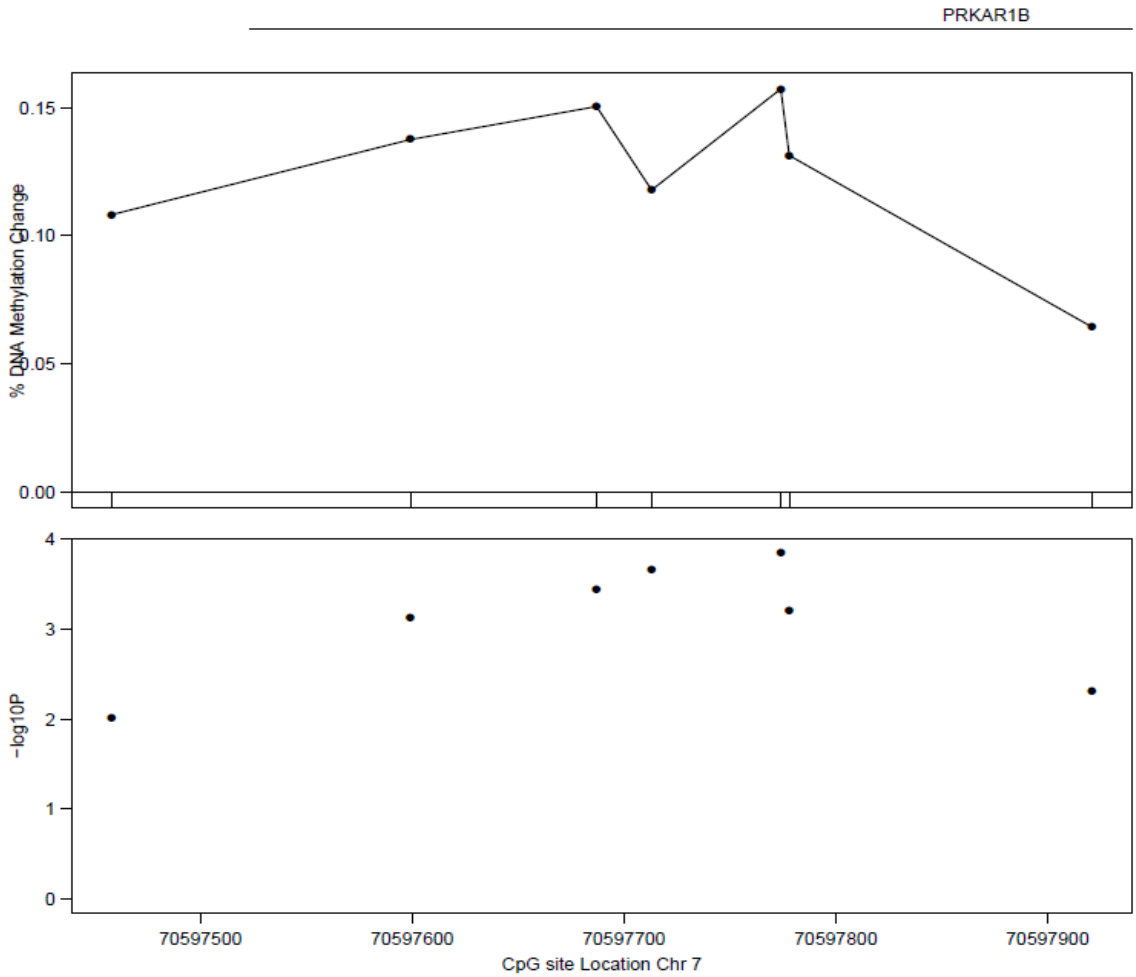


Figure 3.11 The top ranked DMR and annotated gene (PRKAR1B). This graph demonstrates the percentage methylation change of the top ranked gene identified through the DMRs and its distribution on chromosome 7.

Chromosome	Gene	Name	N. Probes	Sidak P Value	CpG Region	Function	Methylated in CRC
chr16:82659960-82660873	CDH13	Cadherin 13	15	5.98E-12	Island	Cell recognition and adhesion, anti-apoptotic. Tumour suppressor gene, involved in cell cycle regulation.	Progression of normal colon to aberrant crypt foci (Section 1.6.3.3)
chr1:25257505-25258082	RUNX3	Runt-related transcription factor 3	13	2.55E-10	Island	Transcription factor	Progression of normal colon to aberrant crypt foci (Section 1.6.3.3)
chr9:41166530-41167278	SFRP1	Secreted frizzled-related protein 1	8	8.79E-16	Island	Wnt antagonist	Progression of normal colon to aberrant crypt foci (Section 1.6.3.3)
chr4:154710224-154710961	SFPR2	Secreted frizzled-related protein 2	17	2.07E-17	Island	Wnt antagonist	Progression of normal colon to aberrant crypt foci (Section 1.6.3.3)
chr4:154710224-154710961	SFPR2	Secreted frizzled-related protein 2	6	4.21E-12	Island	Wnt antagonist	Progression of normal colon to

							aberrant crypt foci (Section 1.6.3.3)
chr6:152128024-152129036	ESR1	Oestrogen receptor 1	21	1.84E-11	Island	Transcription factor	Progression of Aberrant Crypt Focus to Polyp/Adenoma (Section 1.6.3.3)
chr2:182321489-182322841	ITGA4	Integrin Alpha Subunit Gene	11	6.02E-23	Island	Encodes Integrin Alpha 4 receptor for fibronectin	Progression of Aberrant Crypt Focus to Polyp/Adenoma (Section 1.6.3.3)

Table 3.14 The list of genes annotated to DMRs known to play a role in the in the multi-step process from normal colonic epithelium leading to adenocarcinoma.

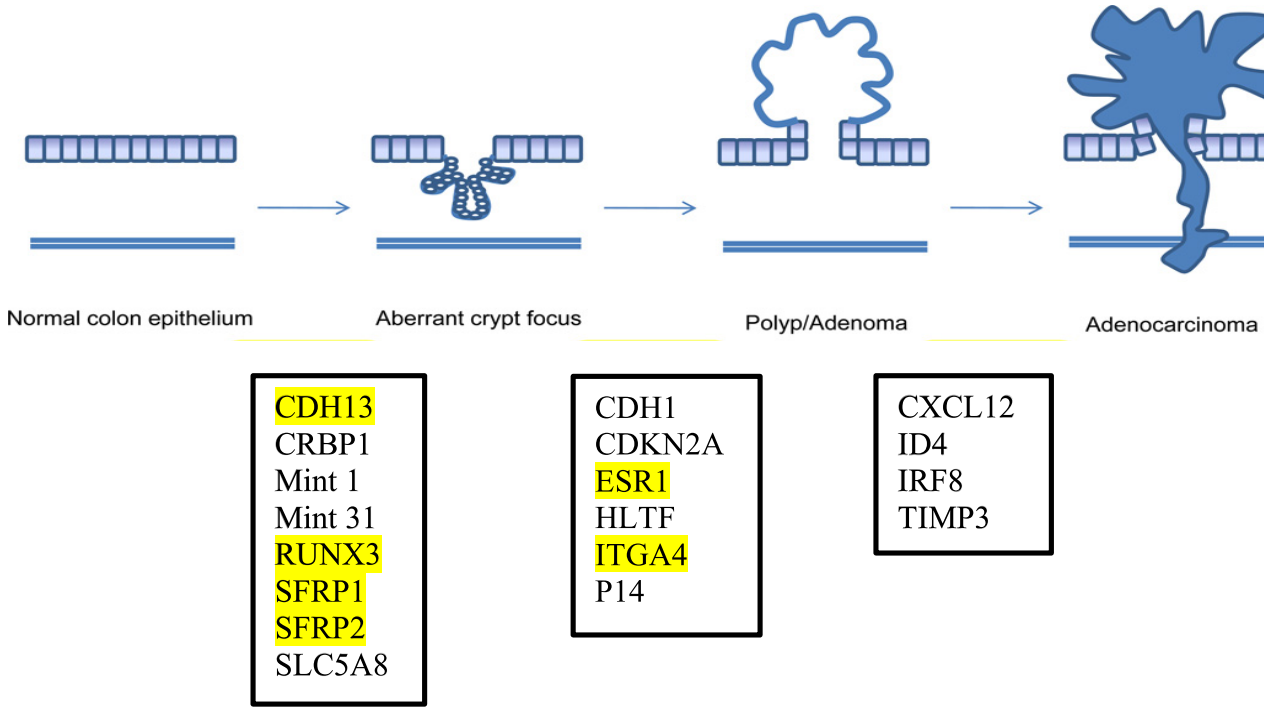


Figure 3.12 The commonly methylated genes in the progression of normal colon to adenocarcinoma with emphasis on rectal cancer associated genes. In black are all genes previously known to play a role in the progression of normal epithelium to adenocarcinoma. Highlighted in yellow are the genes we found hypermethylated in rectal tumour samples compared with normal mucosa. Adapted from Lao and Grady (2011).

Region	Gene	N.Probes	P Value	Epigenetic/Cancer Literature
chr11:128693677-128694679	FLI1	9	7.83E-07	Promising biomarker for rectal cancer as part of a specific DNA methylation signature (Vymetalkova et al., 2016)
chr14:23834710-23835212	EFS	9	6.24E-13	Associated with prediction of biochemical, local, and systemic recurrence of prostate cancer (Vanaja et al., 2009)
chr17:75368750-75369657	Sept9	10	2.21E-05	Used commercially as a methylation biomarker for CRC detection (Payne SR, 2010)
Chr7:19156621-19158747 **Chr7:19146032-19146555**	TWIST1 TWIST1	30 5	4.45E-15 1.94E-09	Differentially methylated in rectal tumours (Exner at al., 2015)
Chr4:111533267-111533951 **Chr4:111554966-111555503** **Chr4:111549880-111550666**	PITX2 PITX2 PITX2	7 5 6	1.41E-10 4.00E-08 2.38E-06	Differentially methylated in rectal tumours (Exner at al., 2015)
Chr11:44330903-44333192	ALX4	46	5.56E-15	Hypermethylated in CRC serum and correlated with advanced stage disease (Ebert et al., 2006)
Chr1:25257505-25258082	RUNX3	13	2.55E-10	Hypermethylation in serum a promising diagnostic marker in CRC (Tan et al., 2007)

Chr8:97505391-97505868 **Chr8:97507561-97507958**	SDC2 SDC2	8 5	1.33E-11 2.08E-11	*SDC2* methylation in serum is useful for diagnosis and monitoring of CRC (Oh, et al., 2013)
Chr4:154710224-154710961 **Chr4 154713235-154713789**	SFRP2 SFRP2	17 6	2.45E-12 6.63E-07	Potential stool and serum biomarker in CRC patients (Nagaska at al., 2009, Tang et al., 2011)
Chr7:93519220-93520566	TFPI2	20	2.34E-20	Hypermethylation is a potential biomarker for detection of CRC (Glockner et al., 2009)
Chr20: 23029640-23030343	THBD	5	4.32E-14	Hypermethylation marker for early CRC detection (Lange et al., 2012)
Chr17:1961109-1961778	HIC1	5	1.21E-12	Hypermethylated in CRC stool samples (Lenhard et al., 2005)
Chr20:58179847-58180616	PHACTR3	5	1.34E-13	Found methylated in CRC stool (Bosch et al., 2012)
Chr13:36919344-36919960	SPG20	7	1.93E-09	Detected in stool samples of patients with CRC (Zhang et al., 2013)

chr9:41166530-41167278	SFRP1	8	8.79E-16	DNA stool test of hypermethylation of SFRP1 promoter found to be a good method for early detection of CRC (Zhang et al., 2007)
Chr4:96470053-96471143	UNC5C	12	1.93E-17	Inactivation of the UNC5C Netrin-1 receptor is associated with tumour progression in colorectal malignancies (Bernet et al., 2007). UNC5C methylation was found to be negatively correlated with protein expression and prognosis in CRC patients (Wu et al., 2017)
Chr4:96468962-96469634	UNC5C	9	2.51E-13	

Table 3.15 The list of genes annotated to DMRs that are known to play a role in rectal cancer or shown to have biomarker potential.

3.3.7 Identifying key pathways: gene ontology

To assess whether the genes identified (Section 3.3.6) were biologically meaningful, GO enrichment analysis was performed on genes annotated to DMPs identified in the analysis using the R package.

The logistic regression approach described by Lunnon et al. (2016) was used for the GO analysis. In summary, all probes were mapped to genes using the standard gene annotation file provided by Illumina (GRCh37/hg19). GO terms were downloaded and annotated to all probes based on these gene annotations where all parent terms were also included. Probes without genic annotation or without any gene ontology annotation were excluded from the subsequent analysis. GO Pathways were filtered to contain between 10 and 2000 genes each. Significant pathways (P <0.05) were then grouped together using a loop algorithm.

We therefore analysed 828 genes annotated with DMRs. Our analysis revealed the significant enrichment of multiple modules in various GO terms. We grouped the GO terms into two different categories: cellular components and biological functions. In the cellular category, our analysis identified 46 enriched (P < 0.05) terms (Table 3.16). The top proteins encoded by the analysed genes in this category, mainly associate with neuronal functions and plasma membranes as expected. Although we are unable to give a full description of each of the significant pathways due to limited space, many are interesting in relation to cancer. For example, the extracellular region (GO: 0005576, p < 0.0013) and its GO parts including the extracellular matrix (ECM) (GO:0031012, p < 7.55151E-08), the proteinaceous ECM (GO:0005578, p < 1.35639E-06), the ECM part (GO:0044420 p < 0.0017) as well as the transcription factor complex pathway (GO:0005667, p < 0.0004) are listed in the top categories and play a role in cancer (Section 4.6).

Categorisation by “biological process” identified 687 enriched (p <0.05) terms, mainly associated with developmental, neuronal and cell differentiation activities. We observed that the top terms included pathways that are known to play a role in cancer, such as DNA cell-cell adhesion (GO:0098609), cell-cell adhesion via plasma-membrane adhesion molecules (GO:0098742), regulation of transcription regulatory region DNA binding (GO:2000677), regulation of transcription regulatory region DNA binding (GO:2000677) and cell fate

commitment (GO:0045165). Table 3.17 lists some of the terms that are relevant to cancer. All the enriched biological and cellular terms can be found in Supplementary Table 5 and Supplementary Table 6 (Appendix 4).

GO	Term	N of Genes	P-value	Adjusted P-value
GO:0044456	Synapse Part	40	1.9776E-09	6.24921E-07
GO:0005887	Integral component of plasma membrane	71	3.36109E-08	5.31052E-06
GO:0031012	Extracellular matrix	34	7.55151E-08	7.95426E-06
GO:0030425	Dendrite	25	1.09876E-06	8.57239E-05
GO:0005578	Proteinaceous extracellular matrix	25	1.35639E-06	8.57239E-05
GO:0034702	Ion channel complex	24	1.38185E-05	0.0007
GO:0045211	Postsynaptic membrane	20	2.01469E-05	0.0009
GO:1902495	Transmembrane transporter complex	25	2.42697E-05	0.0010
GO:1990351	Transporter complex	25	3.16854E-05	0.0011
GO:0097060	Synaptic membrane	21	5.41966E-05	0.0017
GO:0033267	Axon part	15	8.26454E-05	0.0024
GO:0045202	Synapse	21	0.0001	0.0033
GO:0043235	Receptor complex	22	0.0002	0.0051
GO:0005667	Transcription factor complex	21	0.0004	0.0093
GO:0008328	Ionotropic glutamate receptor complex	8	0.0004	0.0093
GO:0030424	Axon	15	0.0006	0.0109
GO:0042734	Presynaptic membrane	7	0.0011	0.0213
GO:0005576	Extracellular region	75	0.0013	0.0223

GO:0044420	Extracellular matrix part	12	0.0017	0.0288
GO:0043025	Neuronal cell body	19	0.0025	0.0402
GO:0005614	Interstitial matrix	4	0.0028	0.0408
GO:0044297	Cell body	21	0.0028	0.0408
GO:0032809	Neuronal cell body membrane	3	0.0045	0.0589
GO:0044298	Cell body membrane	3	0.0045	0.0589
GO:0031045	Dense core granule	3	0.0058	0.0715
GO:0005604	Basement membrane	8	0.0059	0.0715
GO:0043204	Perikaryon	6	0.0074	0.0842
GO:0014069	Postsynaptic density	9	0.0075	0.0842
GO:0043679	Axon terminus	5	0.0078	0.0851
GO:0034707	Chloride channel complex	6	0.0089	0.0912
GO:0030672	Synaptic vesicle membrane	4	0.0089	0.0912
GO:0008021	Synaptic vesicle	8	0.0106	0.1048
GO:0043195	Terminal bouton	5	0.0129	0.1225
GO:0048786	Presynaptic active zone	4	0.0132	0.1225
GO:0005615	Extracellular space	51	0.0154	0.1387
GO:0044306	Neuron projection terminus	5	0.0199	0.1750
GO:0097481	Neuronal postsynaptic density	5	0.0284	0.2429
GO:0031225	Anchored component of membrane	10	0.0337	0.2799
GO:0043197	Dendritic spine	6	0.0371	0.2867
GO:0034705	Potassium channel complex	6	0.0371	0.2867
GO:0005581	Collagen trimer	7	0.0382	0.2867
GO:0061202	Clathrin-sculpted gamma-aminobutyric	2	0.0390	0.2867

	acid transport vesicle membrane			
GO:0032279	Asymmetric synapse	2	0.0390	0.2867
GO:0044309	Neuron spine	6	0.0434	0.3118
GO:0044224	Juxtaparanode Region of axon	2	0.0467	0.3207
GO:0001527	Microfibril	2	0.0467	0.3207

Table 3.16 GO enrichment analysis top categories grouped by cellular function and p value.

	Term	N of genes	P-value	Adjusted P-value
GO:0098742	Cell-cell adhesion via plasma-membrane adhesion molecules	35	4.44464E-14	3.40564E-11
GO:0098609	Cell-cell adhesion	35	5.09521E-14	3.40564E-11
GO:0007156	Homophilic cell adhesion via plasma membrane adhesion molecules	28	3.94219E-12	1.46387E-09
GO:0001708	Cell fate specification	17	8.08533E-10	1.93008E-07
GO:0010721	Negative regulation of cell development	30	1.17763E-08	2.07138E-06
GO:0045165	Cell fate commitment	22	1.6157E-08	2.57127E-06
GO:0001709	Cell fate determination	12	3.11388E-07	3.05004E-05

GO:0007188	Adenylate cyclase-modulating G-protein coupled receptor signalling pathway	19	3.19789E-07	3.05004E-05
GO:0023019	Signal transduction involved in regulation of gene expression	9	6.02994E-07	4.79811E-05
GO:0007193	Adenylate cyclase-inhibiting G-protein coupled receptor signalling pathway	13	1.3342E-06	9.09979E-05
GO:0007187	G-protein coupled receptor signalling pathway, coupled to cyclic nucleotide second messenger	20	1.8803E-06	0.0001
GO:0030855	Epithelial cell differentiation	27	2.54267E-05	0.0011
GO:0090092	Regulation of transmembrane receptor protein serine/threonine kinase signalling pathway	20	2.63132E-05	0.0011
GO:0048545	Response to steroid hormone	29	0.0001	0.0037
GO:0030510	Regulation of BMP signalling pathway	11	0.0001	0.0043
GO:0051101	Regulation of DNA binding	11	0.0002	0.0064
GO:0030178	Negative regulation of Wnt signalling pathway	14	0.0003	0.0090
GO:0043388	Positive regulation of DNA binding	7	0.0004	0.0103

GO:0051954	Positive regulation of amine transport	6	0.0004	0.0104
GO:0030111	Regulation of Wnt signalling pathway	19	0.0004	0.0108
GO:0035235	Ionotropic glutamate receptor signalling pathway	6	0.0005	0.0112
GO:0045785	Positive regulation of cell adhesion	16	0.0007	0.0150
GO:0007263	Nitric oxide mediated signal transduction	5	0.0007	0.0155
GO:0034199	Activation of protein kinase A activity	5	0.0007	0.0155
GO:0090101	Negative regulation of transmembrane receptor protein serine/threonine kinase signalling pathway	12	0.0008	0.0155
GO:0045168	Cell-cell signalling involved in cell fate commitment	6	0.0009	0.0177
GO:0007200	Phospholipase C-activating G-protein coupled receptor signalling pathway	8	0.0009	0.0178
GO:0007215	Glutamate receptor signalling pathway	7	0.0010	0.0187

Table 3.17 GO enrichment analysis cancer related categories grouped by biological function and p value.

Chapter 4 - Discussion

4.1 Introduction

Most studies have considered colonic and rectal cancers as a single entity, despite the molecular, clinical and histopathological differences between them (see Section 1.3.2 and Section 1.4.6). Very few studies to date have looked at DNA methylation of rectal cancer as a separate entity. The majority of those that have, did not take a systematic approach across the whole of the genome and included a limited set of genes. To our knowledge, only two other studies have addressed the genome wide methylation differences between rectal cancer and adjacent normal mucosa.

In this study, we assessed genome wide patterns of DNA methylation in 30 human matched rectal tumour and adjacent normal mucosal samples (from 7 males and 8 females, age range 45-79 years) using the Illumina Infinium 450K array platform.

The first aim of this study was to use and validate methods used in CRC studies, for isolating and processing tumour cells from rectal cancer tissues.

The second aim was to identify methylomic variation associated with rectal cancer across the whole of the genome using the Illumina Infinium 450K array platform, that may form diagnostic of prognostic signatures of primary human rectal carcinoma based on DNA methylation.

The chosen methodology and the lessons learned will be presented first, then the findings will be discussed to shed a light on the potential of DNA methylation in rectal cancer.

4.2 Methodology

Current genome-wide methods to detect DNA-methylation in healthy and diseased tissue require high-quality DNA from FF samples, which provide the gold standard tissue quality (Ruijter et al., 2015). Therefore, the tissues we actively collected as part our study, were swiftly transferred in liquid nitrogen and stored as FF at − 80 °C to minimise nucleic acid degradation. We obtained enough DNA following bisulfite conversion for downstream processing using the Illumina BeadChip 450. Quality control revealed the ‘success’ of the FF tissues as they passed our stringent QC checks (only one sample failed the QC tests

for unidentified reasons) (Section 3.3). Thus, we successfully replicated the methodology commonly used in DNA methylation studies using rectal FF tissues.

We also chose to use historically stored samples, fix them in formalin and process them as FFPE tissues. This is a common clinical method for processing tissues which, unlike the FF method, allows for dissection and quantification of tumour content in biopsied samples.

Whilst FF tissues are the gold standard, FFPE tissues on the other hand are less optimal for DNA methylation studies. First, the fixation process and storage of tissue embedded in paraffin can lead to fragmentation, nucleotide base lesions, modified bases and cross linkage (Iwamoto et al., 1996, Srinivasan et al., 2002). This results in DNA fragmentation and low DNA quality.

Furthermore, the bisulfite modification required for downstream processing using the 450K array, degrades DNA further, making assessment of methylation challenging.

With that in mind, we quantified our DNA (see Section 2.4.4) and modified the protocol (see Section 2.5.1) to account for the challenges of FFPE tissues. Unfortunately, the FFPE samples were suboptimal and did not pass our stringent QC testing. The DNA resulting from our FFPE isolation protocol was too fragmented to be used in any downstream analyses. No current approach (methodological or bioinformatic) would be able to overcome the technical limitations of using these samples. We therefore excluded them our study (see Section 3.2).

We believe that the inherent suboptimal quality of FFPE tissues, the long time spent homogenising the thick fibrous rectal tissue with rotor–stator homogeniser (Section 2.3.1) — which results in greater DNA fragmentation — and the potential initial delays in freezing the biopsied historical samples (2004-2007) may have played roles in the DNA fragmentation of the samples (Section 2.1.3.1).

The poor integrity of FFPE tissues and decreased compatibility with Illumina HumanMethylation 450 BeadChip is recognised. New protocols, technologies and kits are constantly being manufactured to circumvent this problem. Recently, Ruijter el al. (2015), showed that FFPE tissue-derived DNA is successfully restored using the Illumina FFPE DNA Restoration solution kit

(Illumina, San Diego, USA), making it feasible for downstream processing with the 450K array with successful results. Although we were not able to try it as part of the current project, the incorporation of this restorative kit along with the implementation of the Illumina modified FFPE protocol (Illumina, San Diego, USA) are future steps that may ensure better quality DNA from FFPE tissues.

One problem common to both FF and FFPE samples is tissue heterogeneity. CRC tissue is characterised by heterogeneity, cellular and molecular. It combines a mixture of normal epithelial ducts, in situ or invasive tumour cells, adjacent blood vessels, stroma, immune system cells and the extra cellular matrix (ECM) (Bahrami et al., 2017). Subsequently, cancer biopsies will invariably contain dysplastic cells, stromal cells and adjacent morphologically normal cells. Therefore, 'contamination' of the biopsied sample is possible, resulting in either higher or lower quantitative methylation results. Indeed, the results of the histopathological examination of our FFPE tissues showed high variability in tumour content, a discordance between the expected sample type and the actual sample type and sections of dysplasia instead of adenocarcinoma (Section 3.2.1). Manual microdissection can provide a solution to this problem. Isolating tumour cells under microscopy is a relatively easy method of enriching tumour cells by excluding the majority of non-tumour cells and ensuring tumour content is high. However, it does require the time and effort of a histopathologist or a trained person. Laser capture microdissection is also another option which offers higher resolutions and could isolate relatively pure individual cells (Fend and Raffeld, 2000). This technology, however, is expensive and not readily available. Whether such high level of cell purity is necessary for molecular testing in rectal cancer is debatable.

In contrast to FF tissues, FFPE tissues allows for histological assessment. FF tissues however remain the gold standard for DNA methylation. Therefore, we propose a trial of the following methodology for future rectal methylation studies: a) obtaining a core biopsy of the tissue b) dissection of the tissue biopsy into four equal parts c) three parts to be embedded in paraffin and assessed histologically for sample type and tumour content d) the fourth part to remain fresh frozen and used in downstream methylation studies e) estimation of the tumour content of FF section from the histological assessment of the three

FFPE parts. Whilst this is a more complex way of quantifying and qualifying tumour and mucosal tissues for methylation studies in cancer, it offers a potential less expensive and labour intensive solution to the tissue heterogeneity problem whilst allowing the use of the gold standard FF tissues for methylation studies.

In summary, we validated previously used methodology for FF tissues in genome wide and specific DNA methylation gene studies for rectal cancer and proposed a potentially improved methodology for future studies.

4.3 Global methylomic variation in rectal cancer

In this study, we have demonstrated global hypomethylation in rectal tumour samples compared with adjacent normal mucosal samples.

Over three decades ago, the first epigenetic alteration in cancer was described by Feinberg and Vogesltein (1983): global hypomethylation in colon cancer compared with normal colon. This loss of methylation mainly affected repetitive sequences of DNA and satellite repeats and appeared to be an early event in CRC. Since that discovery, CRC remained one of the most studied diseases in the field of epigenetics and DNA methylation research. Less is known however about the global methylation changes specific to rectal cancer, despite the two being increasingly recognised as separate entities.

Global DNA hypomethylation is associated with genomic instability and chromosome abnormalities. It is often seen before the formation of adenomas, can also be observed in the healthy tissue adjacent to tumours suggesting a role in the initiation of disease. Furthermore, hypomethylation is thought to predispose to genomic instability as increasing hypomethylation is observed in the adenoma-carcinoma sequence (Sheaffer et al., 2016).

Global DNA methylation often occurs in repetitive elements of the genome such as long interspersed repeat sequences (LINE-1). Benard et al. (2013) investigated whether global DNA Methylation and histone modification patterns can be used to predict clinical outcome in rectal cancer patients enrolled in a clinical trial. They showed that LINE-1 methylation is an independent predictor of survival and recurrence in early stage rectal cancer.

Contrastingly, a recent study could not identify LINE-1 hypomethylation in the aberrant crypt foci of rectal cancer samples (Quintanilla et al., 2014). Tsang et al. (2014) looked at changes in global DNA methylation levels following treatment in rectal cancer, in a study which included 53 patients who underwent nCRT. They found a correlation between pretreatment DNA methylation levels and treatment response: patients with lower global methylation levels pre nCRT were more likely to achieve pCR. They were also able to predict the likelihood of a pCR and partial pathological response from pre-treatment global methylation level.

We report global DNA hypomethylation in rectal tumour samples compared to normal mucosa. Our findings are in line with previous studies in colon cancer. It is possible that global DNA hypomethylation in the rectum may contribute to tumourigenesis similar to the colon and may play a role in the prediction of treatment response.

4.4 Aberrant DNA methylation in the pathogenesis of rectal cancer

The pathogenesis of rectal cancer is largely linked to that of CRC. The adenoma-carcinoma sequence in CRC is well established. It is characterised by genetic and epigenetic changes in a step wise manner and is responsible for the development of sporadic and familial FAP (Section 1.6.3).
DNA methylation plays a role in CRC and several genes are known to be hypermethylated along different stages of the adenoma-carcinoma progression. We identified DMPs and DMRs annotated to some of these genes, notably: CDH13, RUNX3, SFRP1, SFRP2, ESR1 and ITGA4.
CDH13, RUNX3, SFRP1 and SFRP2 have been shown to be methylated early in the normal colonic epithelium to aberrant crypt focus stage, whereas methylation of ESR1 and ITGA4 is associated with the aberrant crypt focus to adenoma stage. Both CDH13 and ESR1 have previously been found differentially methylated between rectal cancer FFPE tissues and normal mucosa (Leong at al., 2011, Molinari et al., 2013) further confirming our results.

Age may also be a risk factor in the development of CRC and thus in rectal cancer. Aberrant DNA methylation is observed in normal tissues of older individuals. Therefore, CRC in older people may have more epigenetic events than cancers seen in younger people. Approximately half of the genes that show age related methylation are the same genes that are involved in the pathogenesis of colon cancer (Toyota and Issa, 1999).
Some of the genes we identified as aberrantly methylated in rectal cancer have been linked to age associated methylation. Changes in the methylation of ESR1 and TUSC3 for example, occur in histologically normal colonic epithelium in an age-related fashion (Lao and Grady, 2015). ESR1 is often methylated in distal CRC and reported to show age related methylation (Iacopetta, 2002). Belshaw et al. (2008 and 2010) assessed age-dependent gene methylation of a panel of genes which include SFRP1, SFRP2 and ESR1 in normal colon mucosa. They found significant differences in gene methylation levels between people with and without colon cancer.
Not only aberrant DNA methylation in gene promoters is seen in ageing, but global hypomethylation is also seen in the ageing process (Lao and Grady, 2015). This suggests that there may be common mechanisms for age related and cancer related DNA methylation. However, the cause of this age related differential methylation remains unknown.

The concept of field cancerisation has also been investigated in the context of epigenetic events and tumourigenesis in the colon. Studies have shown that epigenetic alterations may be indicators of field cancerisation in the colon, by creating a predisposition for specific cancer related mutations (Ramirez et al., 2008, Worthley et al., 2010). Field cancerisation refers to the susceptibility of normal tissues and cells to develop cancer, possibly because of the environment. It was first proposed to explain the development of multiple primary tumours, local recurrence, abnormal tissue surrounding the cancer and focal areas of precancerous changes (Slaughter et al., 1953). Field cancerisation occurs at the molecular level of precancerous cells, which may have a clonal relationship to the tumour and can persist after complete resection of the tumour. Yet it is still unclear how far the cancer field can extend from the primary tumour.

Inactivation of the APC tumour suppressor gene represents the rate-limiting event in CRC formation. The loss of APC function leads to the constitutive activation of the canonical Wnt signal transduction pathway, which regulates the finely tuned equilibrium between cell proliferation, differentiation and apoptosis. Aberrant Wnt pathway signalling is an early progression event in the majority of CRCs (Fodde et al., 2001, Fodde and Tomlinson, 2010). Interestingly, the hypermethylation of SFRP genes are present in monoclonal aberrant crypt foci lacking APC mutations. This is thought to contribute to constitutive Wnt ligand signalling and decreased apoptosis. The constitutive activation of the pathway results in cellular defect such as proliferation, differentiation, cell migration and genetic instability, leading to the formation of the aberrant crypts - the first stage of the adenoma-carcinoma sequence (Reya and Clevers, 2005). This constitutive activation is required to complement downstream mutations for the evolution of CRC. This epigenetic loss of SFPR function therefore may 'cancerise' the field, rendering cells especially sensitive to further activating events in the Wnt pathway (Suzuki and Bird, 2004). The methylation pattern of SFRP genes in this case may be useful in predicting the risk of developing rectal cancer and further understanding field cancerisation.

We also annotated several other genes to DMPs and DMRs that contribute to the aetiology of other cancers. One of those genes is RNF175, which encodes ring finger protein 175. RING-finger proteins are a large group of proteins concerned with apoptosis, cell-cycle control, and prevention of oncogenesis (Ng et al., 2003). They are also likely to be involved in the ubiquitination pathway, which is thought to contribute to the aetiology of different types of tumours (Hoeller et al., 2006). A recent study by Wang et al (2014) has found that Ring finger proteins, including RNF32, play a role in the progression from Barrett oesophagus to oesophageal carcinoma. Interestingly, RNF14 depletion experiments have demonstrated that it is crucial for colon cancer cell survival through interacting with TCF/β-catenin complexes to regulate Wnt gene transcription (Wu et al., 2013). More information is needed however on the role of ring finger proteins in rectal cancer and their effect on the Wnt pathway.

The relationship between differential methylation in normal rectal mucosa and the susceptibility of the mucosa to undergo cancer formation is complex. Recent

data suggested that hypomethylation of genes that undergo age related DNA methylation correlates with CRCs in CIMP category (Lao and Grady, 2015, Worthley et al., 2010). Whether the differential methylation in normal rectal mucosa is indicative of a field effect or a marker of an associated phenomenon or exposure, such as folate status or tobacco exposure, is still to be determined. One hypothesis suggests that aberrant promoter CpG methylation constitutes an epigenetic field defect that increases the vulnerability of the colon to cancer with advancing age. This theory however remains largely untested.

We have presented in our results several genes involved in the pathogenesis of colon and other cancers. It is likely that many of those play a role in the initiation and progression of rectal cancer, through several pathways and mechanisms, which require further exploration.

4.5 Rectal cancer: CIMP subtype?

Toyota et al. (1999) found two distinct patterns of methylation in CRC. Methylation 'type A', which is age specific and methylation 'type C' which is cancer specific. Type A refers to an increase in global methylation levels with age, whereas type C was called the CIMP: a subgroup of CRC with very high levels of DNA methylation of certain CpG islands, associated with different clinical features, such as age, gender and tumour location, and different molecular features such as BRAF mutation status (Section 1.6.3.2) (Weisenberger et al., 2006). The CIMP is characterised by epigenetic DNA hypermethylation of important cell growth and survival genes, and the subsequent suppression of those genes which are differentially methylated between malignant and normal cells (Williamson et al., 2015).
In the initial discovery, Toyota and ISSA (1999) quantified DNA methylation in 33 cancer genes, using combined bisulfite restriction analysis. They discovered that 7 of those genes were differentially methylated in a subset of cancers known as CIMP +ve. Several alternative methods have been used since for methylation analysis and several methylation panels have been proposed. Weisnberger et al. (20006) contributed to our knowledge by reliably identifying tumours with a CIMP +ve phenotype, consisting of a five-gene panel: CACNA1G, IGF2, NEUROG1, RUNX3 and SOCS1.

Further genes were added in later studies, such as CDKN2A, CRABP1, and promoter region of MGMT, a DNA repair gene (Ogino et al., 2007). In tumours where 4 out of the 5 CIMP panel genes were hypermethylated, a decrease in the expression of p53, p27 and COX2 and an increase in TGFBR2 mutations were found. It was therefore suggested that there are distinct molecular pathways for different CIMP tumours (CIMP high; CIMP intermediate; CIMP low).

This phenotypic classification however faces certain challenges. First, a panel of genes for defining CIMP is yet to be standardised, therefore the definitions of CIMP tumours vary across studies making the generalisation of results difficult. Second, regardless of which panel of CIMP markers is used, CIMP cancers tend to occur in the right colon more than the left, in females more than males and tend to have a high frequency of BRAF mutations (Weisenberger et al., 2006). Additionally, CIMP increases linearly from the rectum to the ascending colon, and is known to be rare in rectal cancer, associated with between 2-10% of cases (Williamson et al., 2015, Jo et al, 2012, Yamauchi et al., 2012). Previous studies looking at CIMP in rectal tumours have reported conflicting results (Jo et al., 2012, Kohonen-Corish et al., 2014., Williamson et al., 2015), which is partly due to a lack of consensus on methods for investigating methylation, the classification into prognostic groups and the utility of these groups in predicting response.

The lack of large cohort size in our study and the lack of prognostic information on patients, some of whom are still undergoing treatment, did not allow us to test for statistically significant methylation signatures for subgroups of our tumour sets at this stage. However, the study is ongoing and long term clinical data is being obtained in order to allow for full CIMP panel testing at a later stage. Meanwhile, in our DMRs list, two differentially methylated annotated genes (RUNX3 and IGF2) which are part of the CIMP panel were identified.

RUNX3 refers to the runt-related transcription factor 3 tumour suppressor gene, located downstream of transforming growth factor-β (TGF-β). This gene regulates the growth and differentiation of epithelial cells and its apoptotic functions are correlated with tumourigenesis and cancer progression (He et al., 2015). RUNX3 protein deficiencies can affect the TGF-β1 signalling pathway by

blocking TGF-β1-induced cell growth inhibition therefore decreasing the sensitivity to apoptosis. Consequently, β-catenin accumulates in the cytoplasm resulting in cell proliferation and apoptosis imbalance. The genetically unstable cell clonal expansion is thought to promote tumourigenesis (Miyazono et al., 2004). Indeed, RUNX3 is involved in the pathogenesis of numerous cancers, such as breast (Chen LF, 2012), ovarian (Lee et al., 2011), head and neck (Tsunematsu et al., 2009) and gastric cancers (Freidrich et al., 2006).
RUNX3 CpG island methylation was also found in colon cancer compared to normal mucosa. Subramaniam et al., (2009) concluded that the inactivation of the gene in colorectal polyps due to promoter hypermethylation is an early event in CRC progression. In a large cohort of both colon and rectal cancers, a strong association between RUNX3 and the risk of colon and rectal cancer was demonstrated (Slattery et al., 2011). Interestingly, in recent a study looking at the effect of RUNX3 on metastasis and angiogenesis in CRC, Kim et al., (2016) found that TNM staging in CRC was correlated with a decreased RUNX3 expression, whereas RUNX3 over-expression inhibited CRC cell migration and invasion. They also demonstrated that the restoration of RUNX3 decreased vascular endothelial growth factor (VEGF) secretion and suppressed tube formation and endothelial cell growth in CRC cells. RUNX3 has also been implicated in recurrence risk in stage II-III colon cancer (Berg et al., 2015). Furthermore, promoter hypermethylation of RUNX3 in serum was found to be a sensitive marker for cancer detection (Tan et al., 2007) showing its promise as a biomarker.

Less is known on the role of hypermethylated IGF2 in CRC. The gene is a member of the insulin/insulin-like growth factor (IGF) signalling pathway. Its receptors are integral membrane proteins and it is involved in the AKT and MAP kinase signalling pathways (Rogers et al., 2016).
Loss of imprinting of IGF2 is an epigenetic change that is known to occur in some cancers. Under normal conditions, only one allele, either maternal or paternal, is expressed. Loss of imprinting of IGF2 refers to the aberrant bi-allelic expression of the gene, resulting in its overexpression and subsequently in pro-mitogenic and growth promoting signals through binding to the insulin-like growth factor receptor type 1 (IGF1R). IGF2 over-expression has been often associated with tumour proliferation, invasion and poor outcome in CRC

(Livingstone, 2013). Rogers et al., (2016) demonstrated an association between IGF2 and proliferation, adhesion and, to a limited extent, apoptosis in CRC cell lines. Similarly, Zhong and colleagues (2017) found that the over-expression of IGF2, a consequence of the hypermethylation of its promoter region, acts as a major tumourigenic driver in a subset of CRC cell lines. Contrary to previous studies, Belharazem et al. (2016) have recently shown that LOI of IGF2 can occur independently of major pathogenic CRC routes (serrated vs. non-serrated), occur throughout the large intestine without predilection for the right colon, and may even constitute a distinct molecular subgroup. Additionally, IGF2 over-expression was associated with treatment resistance in certain CRC subgroups (Zanella et al., 2015) making it an important gene for further studies

In summary, the hypermethylation of RUNX3 and IGF2 has a strong association with the initiation, progression and to a certain extent, treatment response of CRC cancer. Those two genes have mainly been looked at as part of the CIMP panel of CRC subsets. Larger cohorts are required to assess CIMP status in rectal cancer and the subgroups associated with different clinicopathological characteristics if any.

4.6 Biological and cellular pathways in rectal cancer

Our GO analysis revealed several pathways, both in the biological and cellular categories, that are involved in rectal cancer. In the cellular function category, several of the most significant pathways and DMRs annotated to genes, were associated with the ECM functions, such as IGFBP3, MMP13, MMP23B, FBN2, GPC6, COL4A2, NTNG1, ITGA4 and SPOCK1 amongst others (Table 3.16) (Simmer et al, 2012).

The ECM is synonymous with the proteinaceous ECM and is part of the extracellular region (Section 3.3.7). Figure (3.13) represents this relationship through an ancestor chart. It is an elaborate meshwork of cross-linked proteins that provide architectural and structural support for cells. These proteins bind to growth factors, present them to cells and provide biochemical cues that are major regulators of cellular behaviour (Hynes, 2009). The ECM affects motility,

viability, adhesion and proliferation of cells and constitutes a major part of their environment. Subsequently, it is also a major component of a tumour's microenvironment: it has the potential to support its survival and proliferation, sustain their growth, encourage the tumour cell's seeding and invasion of nearby as well as distant sites (Cretu and Brooks, 2007). Therefore, deregulation of ECM pathways may play a role in the pathogenesis and progression of tumour cells.

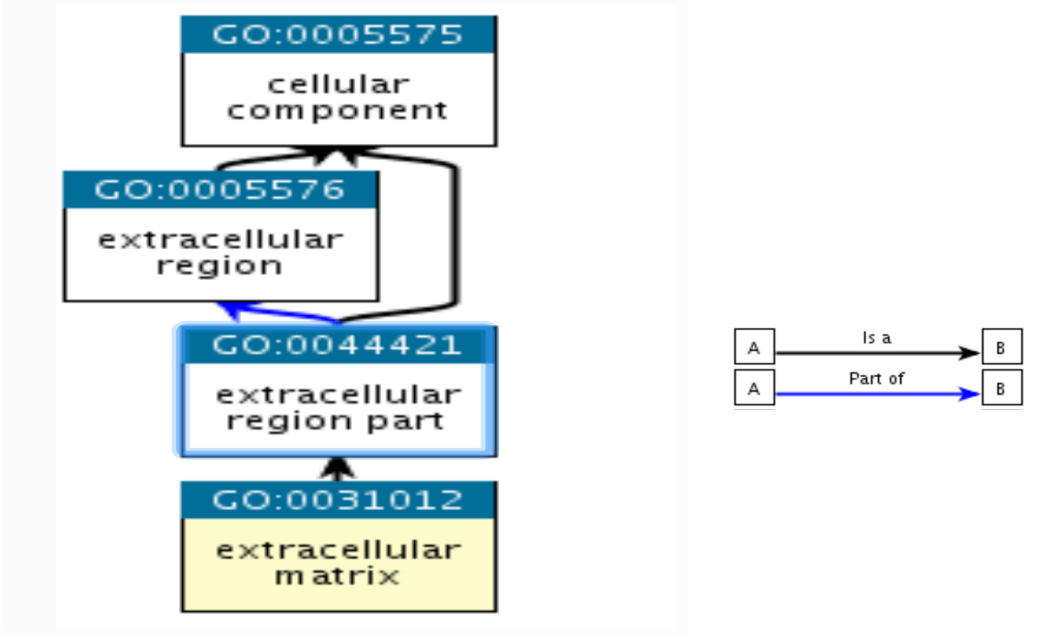


Figure 3.13. Ancestor chart for the ECM (GO:0031012). Adapted from the European Bioinformatics Institute (2018).

Aberrant ECM environments have also been implicated in the ability of tumours to undergo epithelial mesenchymal transition (EMT) and metastasise. This is because cells lose cell-cell and cell-extracellular matrix adherence along with profound changes in their cytoskeleton architecture. DNA methylation induces silencing of ECM pathway regulatory genes (Yi et al., 2011). A recent study found that the methylation of ITGA4 and TFPI2 is an early and frequent event in precancerous and cancerous lesions of the rectum (Gerecke et al., 2015). The TFPI2 gene encodes a serine proteinase inhibitor that protects the ECM of cancer cells from degradation and inhibits in vitro colony formation and proliferation. Hypermethylation induced loss of TFPI2 function could

predispose cells toward a pro-invasive program in rectal cancer, an important event especially in later stages of carcinogenesis (Glockner et al. 2009, Hahn et al. 2008).
The ITGA4 gene encodes the alpha-4 subunit of an integrin family member. Integrins are membranous receptor proteins which play important roles in extracellular control of cell survival and differentiation via cell–cell interactions and cell–matrix communication. ITGA4 therefore is able to control cell adhesion, making it a putative tumour suppressor (Ausch et al. 2009). A loss in function is shown to be an early event in carcinogenesis through aberrations of the ECM and cell-cell pathway integrity.

Interestingly, it is also thought that the ECM has a potential to confer treatment resistance, by acting as a barrier to drug delivery. In a recent study on CRC cell lines, Stankevisius et al., (2016) established ECM-dependent pathways which may be feasible targets for anti-cancer therapies. Other studies however, demonstrated that inhibition of the stroma reorganisation and alignment through drugs or genetic engineering, accelerates tumour growth and decreases survival (Ozdemir et al., 2014; Rhim et al., 2014). This suggests that is a more complex role for the ECM in cancer homeostasis and progression.

The idea that changes in ECM architecture may allow cancer development and progression is novel, and the possibility that these changes could be used as diagnostic or therapeutic tools in cancer treatment is still poorly investigated. We have demonstrated an association between deregulated ECM pathways though aberrant methylation of genes in rectal cancer. Further research is required to establish the role of ECM in rectal cancer and its potential as a therapeutic tool.
We also observed that some of the top terms included in our GO biological analysis were related to cell-cell adhesion, plasma-membrane adhesion molecules regulation of transcription regulatory region DNA binding and cell fate commitment. This is consistent with the knowledge that defective function of cell cycle and cell proliferation regulators is a main cause for tumour development and progression. Cell adhesion molecules play a role in the progression of cancer and interactions of cancer cells with the endothelium determine the metastatic spread of tumours. Additionally, direct tumour cell interactions with

platelets, leukocytes, extra and intracellular components contribute to cancer cell adhesion, extravasation and metastases (Bendas and Borsig, 2011).

A hallmark of cancer is the alteration of cellular differentiation pathways during developmental processes. Mutations in the WNT pathway cause colon cancer through constitutive activation of the nuclear β-catenin/TCF transcription factor complex, identified in GO analysis. Many of the DMPs identified in this study are annotated to genes involved in the Wnt signalling pathway, including SFRP1, SFRP2 and DKK3. Recent research indicates that the Wnt and bone morphogenetic proteins (BMP) pathways, both of which were included in our top results, play a role in normal epithelial cell differentiation, and, if altered, may contribute to CRC tumourigenesis (Bertrand et al., 2012). The Wnt family consists of 19 glycoproteins that have been implicated in diverse biological processes, including cell fate specification, cell proliferation, cell migration, dorsal axis formation and asymmetric cell division. Wnt signalling is activated at the bottom of the intestinal crypts, which is crucial for stem cell maintenance and tissue homeostasis. Aberrant Wnt activation is frequently observed in CRC (Cancer Genome Atlas Network, 2012). The Wnt pathway is commonly divided into β-catenin dependent (canonical) and independent (non-canonical) signalling (Zhan et al., 2016). In the canonical pathway, Wnt ligands binds to the Frizzled receptor and the LRP5/6 co-receptor. This stabilises the gene and leads to cytoplasmic accumulation and nuclear translocation of β-catenin. Stabilised, free β-catenin that translocates to the nucleus associates with TCF/LEF to regulate transcription. The involvement of β-catenin defines the canonical Wnt signalling, compared with the Wnt/Frizzled signalling pathways which do not include β-catenin (Bertrand et al., 2012).
Epigenetic silencing of Wnt inhibitors by DNA hypermethylation has been suggested as a mechanism for the aberrant activation of the Wnt pathway in CRC. Indeed, the epigenetic silencing of extracellular Wnt inhibitors SFRP1, SFRP2 and DKK3, have been reported in several studies (Qi et al., 2006, Rawson et al., 2011). The effect of aberrant canonical Wnt signalling is not only restricted to cancer cells, but can also dynamically interact with the microenvironment of cells and immune system. The non-canonical Wnt signalling also plays a role in the progression of tumours as it can mediate motility of cancer cells during metastasis (Bertrand et al., 2012). From a clinical

perspective, there are important implications to understanding the Wnt signalling mechanisms and role in rectal cancer. NSAIDs such as aspirin, indomethacin and celecoxib, have been shown to have some effect in inhibiting Wnt signalling (Gurpinar et al., 2013). These Cox-2 inhibitors are thought to inhibit Wnt by suppressing prostaglandin E2 (PGE2) production, since PGE2 has been shown to promote Wnt signalling. This may prove valuable in the clinical treatment of rectal cancer and in identifying protective patient factors such as regular medications.

BMPs are members of the TGF β superfamily, that bind to serine-threonine kinase receptors (the serine/threonine kinase signalling pathway was also included in our gene ontology findings) resulting in the activation and nuclear localisation of Smad4 (Bertrand et al., 2012). The loss of BMP signalling is associated with sporadic colon cancers and correlates with TRG. It is also thought to contribute to tumour progression and has been reported to promote the growth of colon carcinomas. Interestingly, there is increasing evidence that mutations affecting BMP pathways corroborate with activated Wnt to drive sporadic colon cancers, particularly in later stages of the disease (Hardwick et al., 2008). Given the importance of these pathways in CRC and the potential for the development of pathway inhibitory mechanisms, they should be examined further to determine their role and therapeutic utility in rectal cancer.

4.7 Concordance with previous rectal cancer studies

In a recent epigenome wide analysis of 25 pairs of rectal cancer vs normal mucosa, Vymetalkova et al. (2016) used the Illumina Human Methylation 450 BeadChip to assess DNA methylation. They identified the genes SND1, ITGA4, OPLAH, NPY, FLI1, TFPI2, ADHFE1, PRKAR1B, TRBJ2–6 annotated to their top ten hypermethylated CpG sites, and other differentially methylated genes including GPR85, GPR88, CHST2 and TLX1, associated with rectal cancer. Remarkably, our results included all of these genes- with the exception of TRBJ2-6- therefore supporting their findings.

Some of these genes are known to play a role in cancer, including SND1 ADHFE1 and OPLAH which have been selected as part of a robustly methylated panel of genes in CRC and cross validated using data from the Cancer Genome Atlas (Naumov et al., 2013).

Exploring these genes further, we learned that OPLAH 3′ UTR hypermethylation has been reported as a shared feature in some tumours (Xu et al., 2012) whilst the promoter hypermethylation of ADHFE1 is frequently present in CRC. A recent study found that alcohol induces methylation-mediated down expression of ADHFE1 and proliferation of CRC cells (Moon et al., 2014). Alcohol induced methylation changes could therefore be interesting to explore. The Staphylococcal nuclease homology domain 1 (SND1) is over-expressed in breast, prostate, colorectal and hepatocellular carcinomas (Jariwala et al., 2015). Over-expression of SND1 is thought to occur at a very early stage in colon carcinogenesis and contribute to the regulation of key players in CRC including APC and β-catenin (Tsuchiya et al., 2007).

Another important gene is PRKAR1B which appeared twice in our list of top ten DMPs, yet little is documented on the role of the gene in CRC. The protein encoded by PRKAR1B is a regulatory subunit of cyclic AMP dependent protein kinase A (PKA), which is involved in the signalling pathway of the second messenger cAMP. Studies have revealed that cAMP/PKA signalling pathway is altered in different cancers and may be used in cancer therapy and diagnosis (Naviglio et al., 2009). Further investigation of aberrant pathways induced by epigenetic changes in this gene may be rewarding.

In another genome wide study, Wei et al. (2016) found 36 genes differentially methylated in rectal cancer, notably the gene GFRA1, which was also annotated to our list of top DMRs. They performed further analysis on this gene and demonstrated its suitability as a potential diagnostic biomarker for rectal cancer.

Leong et al. (2011) in another study evaluated DNA methylation in sporadic rectal cancers treated by Radical TME to identify markers associated with local and distant metastasis. Methylation levels in cancer were analysed by MS-MLOA and compared with methylation levels in adjacent tissue. 24 frequently methylated tumour suppressor genes were looked at. Five were found to be differentially methylated in cancer compared with normal controls. Two of those

genes, ESR1, CDH13, were included in our findings, further highlighting their presence in rectal cancer.

In conclusion, we found large agreement between our results and previously published rectal cancer epigenetic studies, providing further validity to our findings.

4.8 Clinical biomarkers for rectal cancer

The discovery of clinical biomarkers for rectal cancer has been hampered by the lack of research addressing rectal cancer as a separate entity from colon cancer and the use of limited sets of genes in the majority of studies. We identified in our results several DMRs and DMPs annotated to genes that have promising potential as biomarkers. Some may be used in screening, staging and prognosis whilst others in predicting the response to therapeutic intervention. Ultimately the identification of DNA methylation signatures in easily-accessible peripheral tissues such as whole blood would be optimal for use as a diagnostic and prognostic biomarker. Future studies should aim to profile matched blood samples from the same individuals from whom mucosal/tumour samples have been assessed to explore the extent of methylomic covariation between tissues.

Screening, staging and prognosis

Two main methods are currently used for CRC screening: colonoscopy and faecal occult blood test (FOBT). Colonoscopy is the most accurate, however it is costly and is associated with poor patient compliance and procedure related complications. FOBT is a less costly and simple method, but it has low specificity and sensitivity (Smith et al., 2017). Accurate methods for early detection, staging and prognosis of primary and recurrent cancer are extremely important for increasing patient survival.

A screening test for CRC performed via a blood sample is considered an inexpensive simple test that would encourage more patients to undergo screening. It could ultimately result in decreased mortality, cost savings to healthcare since more CRCs would be detected at an earlier stage (Warren et al., 2011). Lofton-Day et al. (2008) described molecular markers for CRC that

are shed from solid tumours into the blood stream. One of these markers is SEPT9, a differentially methylated gene in CRC tissues, which can be sensitively and specifically detected in blood plasma. The hypermethylated SEPT9 was annotated to one of our DMRs in rectal cancer vs normal mucosa. SEPT9 gene methylation has been implicated in CRC for over a decade, and used as a biomarker for over 8 years. In fact, the aberrantly methylated Sept9 in plasma has shown almost 70% sensitivity and 90% specificity for detection of CRC (Bingsheng Li ref 2016). A recent systematic review of the performance of SEPT9 gene methylation assay in CRC concluded that SEPT9 assay is satisfactory for the diagnosis and screening of CRC, with more research needed for treatment monitoring and prognosis prediction. In Europe and the Middle East, a blood based assay that detects methylated SEPT9 is already marketed as a colon cancer screening assay under the name Epi *pro*Colon (Song L, 2017). Whether SEPT9 will achieve similar promising results in rectal cancer specifically is yet to be tested, however the identification of this gene in our results shows a promising first step.

Other aberrantly methylated genes have also shown promising results as future biomarkers. In a recent study assessing the potential of methylated DNA as biomarkers in CRC, Mitchell et al., (2014) selected 23 hypermethylated genes to assess. 11 of those genes had methylation levels detectable in white blood cell DNA therefore making them suitable for further evaluation as blood-based CRC diagnostic biomarkers. We described 7 of them in our rectal tumour DMRs: Col4A2, SDC2, DLX5, FGF5, FOXF1, IRF4, SOX21. The same study identified 6 genes that they considered candidate biomarkers for stool based assays. We identified 3 of them in our DMRs: SL6A15, SOX21 and NPY.
Of note is the IRF4 gene, which belongs to the Interferon Regulatory Factor (IRF) family of transcription factors. the IRF family contributes to the regulation of the interferon signalling pathway (IFN), which has been found to play a role in rectal cancer (Slattery et al., 2011).
Another potential biomarker identified in our results is SDC2, which is a cell surface receptor that appears to play a conflicting role in cancer literature: SDC2 over-expression has a poor prognostic factor in prostate cancer (Popović et al., 2010) whereas its downregulation was associated with poor prognosis in oesophageal squamous cell carcinoma (Huang et al., 2009). In addition to its

recognition in Mitchell's study (2014), *SDC2* was ranked second by Simmer and colleagues (2012) in a survey of genes methylated in CRC and its potential as a plasma biomarker was also demonstrated by Oh et al. (2013).

Several other genes amongst our findings could act as potential biomarkers for rectal cancer screening, staging and prognosis, such as: AQP5, which is associated with metastasis in colon cancer (Kang et al., 2008) and with tumour TNM stage, lymph node and distant metastases in CRC (Shan et al., 2014); methylated C9orf50 which has been identified as a novel blood based DNA methylation biomarker for CRC early detection; FGF12, of the fibroblast growth factor (FGF) family, which is involved in broad mitogenic and cell survival activities and has recently been identified as a novel hypermethylated gene in CRC (Li et al., 2012); IKZF1 which is found to be methylated in several cancer types and has recently shown potential to be used as a blood biomarker in the detection of CRC recurrence (Pedersen et al., 2015); PHACTR3, a potentially suitable biomarker for CRC detection in stool (Bosch et al., 2012); EFS gene, associated with prediction of biochemical, local, and systemic recurrence of prostate cancer (Vanaja et al., 2009) and may be a promising biomarker for rectal cancer as part of a specific DNA methylation signature (Vymetalkova et al., 2016); SFRP2 which has shown good potential for detection as a stool and serum biomarker in CRC patients (Nagaska at al., 2009, Tang et al., 2011); HIC1, a growth regulatory and tumour suppressor gene associated with CRC and detected in CRC stool samples in a hypermethylated state (Lenhard et al., 2005); IRAK3, found to be frequently methylated in hepatocellular carcinoma tissues and associated with tumour stage and poor prognosis (Kuo et al., 2015); CDO1 methylated gene, which has recently been considered to play a role in the CRC and studies have shown it to be a promising biomarker for CRC detection (Yamashita et al., 2014, Vedeld et al., 2015) and the ITGA4 and GATA4 which were included in a recent systematic review looking at DNA based biomarkers in CRC detection and name as some of the most important, noteworthy and independently validated genes to study (Lam et al., 2016)

Treatment response

Response to therapeutic treatment in rectal cancer varies. Up to 15% of patients - in selected series of advanced disease in the UK and across cancer

networks- achieve a cCR (and radiological response) and another 10% of patients who undergo resection, achieve a pCR) with no evidence of tumour remaining in the specimen (Dalton et al., 2012, Renehan et al., 2016). These two groups make up to 25% of rectal cancer patients selected for CRT based upon MRI-stage. Patients who achieve complete response have a significantly improved 5-year survival (up to 85%-100%) compared with patients who show minimal response (55%-66%) (Park et al., 2011, Yeo et al., 2010).

In many centres, without evidence of CRM involvement, CRT is avoided for early stage rectal tumours and surgery is offered as the sole treatment. This contrasts with the Habr-Gama group (2004) who offered CRT to all low rectal tumours and achieved higher rates of cCR.

The molecular heterogeneity of colorectal and therefore rectal cancer is believed to be one of the factors responsible for the variability in treatment response among patients with the same stage of cancer (Buczacki and Davies, 2014). Under the TNM staging (Section 1.2.4 and Section 1.4.4) and treatment guidelines, there is both over and under-treatment of rectal cancer patients. TME for early stage rectal cancer carries with it significant side effects, including bladder and bowel dysfunction. Up to 76% of patients undergoing sphincter preserving surgery will experience changes in bowel habit, the so called low anterior resection syndrome (LARS), which can last for months to years (Dulskas et al., 2016) and has a significant detrimental effect on patient quality of life. Preoperative RT is an additional risk factor for LARS. It also entails important morbidity related to the increased risk of anastomotic leak, tenesmus and neurological toxicity amongst others (Biondo et al., 2016). In addition, preoperative chemotherapy in advanced stage rectal cancer is associated with significant adverse effects, including faecal and urinary incontinence, delayed wound healing and sexual dysfunction (Williamson et al., 2015) as well as the specific side effects of the chemotherapy agents chosen in treatment.

The challenge therefore is to identify patients who are likely to achieve complete response and spare them the morbidity of unnecessary treatment. Subsequently, an accurate method for predicting the response to CRT would allow for improved treatment choice with avoidance of unnecessary side effects: patients with chemoradiation-resistant disease would be spared the morbidity of this treatment. Equally, patients with a prediction of pCR may be spared radical surgery and its complications.

Studies into epigenetic silencing through aberrant DNA methylation have revealed novel insights into the response of tumours to chemotherapy and RT in a number of cancers including cervical, prostate, lung, oesophageal and CRC (Williamson et al., 2015). There are currently no available biomarkers for assessing treatment response in rectal cancer patients. We will discuss some of the interesting genes we identified in our results which have the potential to be investigated as therapeutic response biomarkers.

Exner and colleagues (2015) have assessed DNA methylation differences in rectal cancer using targeted DNA microarray for 360 marker candidate genes. The 'signature' genes that differentiated tumours from adjacent normal mucosa comprised TWIST1 (Twist basic helix-loop-helix transcription factor 1) and PITX2 (paired like homeodomain 2). Both TWIST1 and PITX2 are hypermethylated in different cancers and have been suggested as potential biomarkers. PITX2 methylation has been well established as a prognostic biomarker for breast and prostate cancer (Mikeska et al., 2012). DNA methylation of PITX2 was also associated with prognosis for OS of patients with biliary tract cancers (Uhl et al., 2016). High expression of TWIST1 has pointed to high rates of metastasis, poor prognosis, EMT (a crucial event for dissemination of epithelial tumours) and chemotherapeutic resistance (Brabletz, 2012, Eide et al., 2013, Qin et al., 2012). The potential for TWIST1 as a therapeutic target has recently been investigated. Studies showed that Doxycyline, Sirtuin SIRT6 and RNA-based inhibition of TWIST1 can target the gene in lung and breast cancer cells therefore reducing metastatic potential (Qin et al., 2015, Finlay et al., 2015, Han et al., 2014).
Another important gene is TFPI2 (tissue factor pathway inhibitor 2), which belongs to a family of Kunitz type serine protease inhibitors. TFPI2 is implicated in tissue remodelling, is known to protect the ECM of cancer cells from degradation and to inhibit in vitro colony formation and proliferation (Wong et al., 2007). As discussed earlier, proteolytic degradation of the ECM is considered to be an essential step in tumour growth and metastasis. The loss of the TFPI2 function is thought to negatively regulate the activity of ECM associated enzymes that protect the ECM stability, therefore playing a role in carcinogenesis and its progression (Rao et al., 1998 and 2000).

As such, TFPI2 may act as a tumour suppressor gene and has been found methylated in different tumour types, including pancreatic, cervical, gastric and prostate carcinoma (Exner et al., 2015). Hibi et al. (2011) showed that in TFPI2 plasma hypermethylation was present in 38 CRC patients preoperatively. The Hypermethylation in TFPI2 however was absent in patient who were curatively resected. This suggests that TFPI2 may act as a marker for complete response following treatment. Furthermore, Glockner et al. (2005) demonstrated the potential of TFPI2 as a biomarker by detecting TFPI2 methylation in stool DNA. TFPI2 was also suggested as a potential serum methylation biomarker in rectal cancer (Exner et al., 2015)

Two other genes may be of interest in treatment response studies. First, the TFAP2A gene which is part of the AP2 transcription factor family. The TFAP2A family consists of five subtypes: AP2- α, β, χ, δ and ε. TFAP2A has been shown to modulate the Wnt signalling pathway and the APC protein, key pathways in CRC (Beggs et al., 2015). Interestingly, TFAP2E -a subtype of the AP2 transcription factor family- was associated with chemoresistance to 5-fluorouracil in CRC (Ebert et al., 2012). Beggs et al. (2015) investigated this further, and concluded that it does play a role in chemoresistance but through more complex pathways than previously thought. The second gene is CYP1B1, a cytochrome P450 enzyme which is consistently over-expressed in certain cancers. CYP1B1 has been shown to be active within tumours and possesses the ability to metabolise a diverse range of anticancer drugs. CpG methylation of the CYP1B1 promoter region epigenetically regulates CYP1B1 expression in cancers including CRCs. Cancers with aberrant CYP1B1 expression might show altered response to chemotherapy in CRC (Habano et al., 2009) whilst in epithelial ovarian cancer cells, CYP1B1 enhanced the resistance of the cells to chemotherapy (Zhu et al., 2015). Interestingly, CYP1B1 methylation in the bone marrow of acute myeloid leukaemia patients changed during their chemotherapy treatment (Xia et al., 2016).

We have included patients in our study that underwent nCRT with the aim of reinvestigating our findings in the future and looking at potential biomarkers for rectal cancer staging, prognosis and treatment response. We believe this is important for the continuation of this study, as available tissue for the

assessment of biomarkers in rectal cancer might be possible only after the neoadjuvant treatment.

4.9 Summary

In this study, we assessed genome wide patterns of DNA methylation in 30 matched rectal tumour and adjacent normal mucosal samples using the Illumina Infinium 450K array platform. We found that there is global DNA hypomethylation associated with rectal cancer. We identified DMPs and DMRs annotated to genes associated with rectal tumours. Some of the genes we found are associated with the initiation and progression of the adenoma-carcinoma sequence in CRC, such as CDH13, RUNX3, SFRP1, SFRP2, ESR1 and ITGA4. Due to the limited size of our cohort and lack of prognostic information, we did not assess CIMP status in rectal tumours, but identified two genes differentially methylated in our results, RUNX3 and IGF2 which are part of the CIMP panel. We found good agreement with the few previous studies looking at genomic wide variations in rectal cancer with the panels of differentially methylated genes and implicated pathways. The pathways implicated in rectal cancer included Wnt signalling, BMP pathways and ECM involved pathways, in line with CRC aberrant pathways. We also found several genes that may be promising as biomarkers for detection, staging and prognosis of rectal cancer.

4.10 Limitations

Our study has some important limitations. We acknowledge that the number of samples profiled in this study is small, therefore replication in similar larger cohorts is required.

We also recognise that it is a single institution study, of patients from or living in the Devon area in England, as such our sample is not fully representative. Additionally, we included all MRI staged clinical diagnosis of rectal cancer as histological samples taken were considered too superficial to be representative. A next step would be to segregate the rectal tissues into separate categories

(tumour, dysplasia, aberrant crypts) and comparatively assess the methylation levels at each step of the adenoma-carcinoma sequence.

We performed stringent quality control on our processed samples using the 450K array. Whilst the Illumina Human Methylation 450 BeadChip is one of the more advanced options for genome-wide DNA methylation analyses, there are some issues to consider when using this array. First, it interrogates approximately 1.7% of all CpG sites in the human genome (482 422 cytosines) which is substantially less than other methods. Second, the technology has preselected probes that only interrogate specific CpG sites previously identified in methylation-based assays, making the design biased. Third, it works on a co-methylation assumption, which is the assumption that CpG sites located next to those interrogated by the probes will be similarly methylated or unmethylated.

We discussed the methodology adopted in our study - which is a common methodology for DNA methylation studies. We cannot rule out that the observed differences in DNA methylation were caused by carcinogenesis in tumour tissue, or by other confounding factors, such as age, environmental exposures or risk factors. Furthermore, we recognise that the mechanistic relevance of the differentially methylated genes in CRC as noted in our literature review, has not been demonstrated in all instances. It is feasible that many of the differences do not have a causal role in the oncogenetic process but are a consequence of the cancer.

Yet, the site of effects and concordance with previous studies suggest that our methodology and findings are robust. Additionally, the study was robustly done, using stringent quality control measures and advanced statistical analyses. Importantly, it represents a unique experiment cohort, as it is one of the few studies to look at genome wide differential methylation between normal and tumour tissue in rectal cancer as a single entity, separate from the colon. Despite the limitations encountered, our data is consistent with previous findings in CRC and provides further support for the role of DNA methylation in rectal cancer.

4.11 Conclusion

To this date, very few studies have looked at the genome wide changes in rectal cancer as a single entity using the latest sequencing technologies. Our study focused on the discovery of *de novo* epigenetic changes associated with rectal cancer. Epigenetic instability in rectal cancer appears to be manifested in global DNA hypomethylation and gene promoter hypermethylation in CpG islands, similar to CRC. We identified several novel DMPs, DMRs and molecular pathways differentiating rectal tumours from normal mucosa.
The findings of this study are in line with the growing appreciation that aberrant epigenetic events are important in the process of cancer initiation, progression and response to treatment. Larger studies are now required to replicate these results.

Future studies looking at the epigenetics of rectal cancer should assess genome wide and CpG site methylation across the different stages of rectal cancer, from aberrant crypt foci to advanced stage adenocarcinoma, address the relationship of age and DNA methylation, field cancerisation, environmental factors, perform network and pathway analyses and identify potential subgroups as well as biomarkers specific to the disease.

Appendix 1 – Good Clinical Practice Certificate

NHS
National Institute for Health Research
Clinical Research Network

CERTIFICATE of ACHIEVEMENT

This is to certify that

Rachel Dbeis

has completed the course

Introduction to Good Clinical Practice eLearning (Secondary Care)

November 10, 2015

Modules completed:

Introduction to Research in the NHS
Good Clinical Practice and Standards in Research
Study Set Up and Responsibilities
The Process of Informed Consent
Data Collection and Documentation
Safety Reporting

This course is worth 4 CPD credits

CPD
The CPD Certification Service

National Institute for Health Research
Clinical Research Network

Appendix 2 – Application for samples from the tissue bank

Application for Samples from the Royal Devon & Exeter Tissue Bank

Your details

Name of lead researcher: Professor Jon Mills

Contact email:

Contact telephone:

Sample Details

Type of sample required:

- Matched benign and malignant rectal cancer samples

Proposed analysis:

- DNA methylation studies

Numbers of samples and volumes required:
(please provide details of power calculation overleaf)

30 matched benign and malignant samples (60 in total)

(see attached protocol for sample size details)

Did you specifically request the RD&E Tissue Bank to collect these tissues?

NO

If YES please provide the CTB and STB codes allocated to the collection *(found on steering committee approval letter)*:

Do you require any of the following? *(some of these services may be chargeable):*
Matched benign and malignant tissue X
Extracted DNA ☐
Extracted RNA ☐

Accurate histopathological analysis of adjacent tissue ☐
Peripheral blood samples ☐

Do you have prior ethical approval for this study? No
If YES please state ethics number:
If you have ethical approval from a REC to use these samples please provide a copy of your IRAS form and approval letter.

Please provide the following information:
See attached protocol for details requested below

Lay Summary *(explain in less than 300 words why this collection is important and how it will contribute to future improvements in patient care.)*

Scientific Justification *(please provide a full scientific protocol justifying your research methods and outcome measures.)*

Statistics *(please provide an explanation to justify the volumes and numbers of samples requested.)*

Funding and Expertise *(please provide evidence of your expertise to undertake this project (CV) and details of where you expect to fund your analysis from.)*

Signature of Lead Researcher

Date of application:

Please submit completed application to:
rde-tr.tissuebanks@nhs.net

Appendix 3 - Proposal template for use of samples from the RDTEB

JMRD/Vs1/03/03/16

Title: Refining methods for obtaining rectal tumour cells for DNA methylation profiling in rectal cancer biopsies.

Project team:
Professor Jon Mill, Professor of Epigenetics, University of Exeter Medical School; J.Mill@exeter.ac.uk
Mr Neil J Smart, Consultant Colorectal Surgeon RD&E NHS FT
Dr Ian Chandler, Consultant Pathologist RD&E NHS FT
Miss Rachel Dbeis, MRes Intercalating Medical Student
Dr Bridget Knight, Nurse Manger, Exeter Tissue Bank/NIHR Exeter CRF

Summary:

There are approximately 15,000 new cases of rectal cancer each year in the UK. Surgery is the main treatment in early stage cancer however it carries significant risks and intermediate probability of recurrence rates in most patients. Tremendous advances have been made in understanding the molecular pathology of cancer, including rectal cancer. Recent focus has been on DNA sequence mutations and epigenetic variation in tumour cells, although work in this area is still in its infancy.

Chemotherapy and radiotherapy are recommended prior to surgery in advanced stage rectal cancer[1]. Patients' response to treatment is varied: some patients achieve complete clinical response while others do not. Chemotherapy and radiotherapy are both associated with significant adverse effects. Being able to predict which patients are likely to respond to therapy, or have a higher recurrence rate following surgery will spare patients avoidable morbidities, improve care and clinical decision making.

Response to treatment, recurrence rates and prognosis may be dependent on many factors. Recent interest has focussed on the potential role of epigenetic variation in mediating underlying differences in treatment response given the fact that epigenetic dysfunction is known to play an important role in the aetiology, progression and metastasis of cancer[2].

A few studies have started to examine epigenetic variation in rectal cancer, yet a solid methodology for DNA methylation analysis in tumour cells obtained from patients is still to be defined. Of note, no study has aimed to identify genome-wide patterns of differential methylation in rectal cancer, despite the successes of this approach in other types of cancer.

This project aims to develop and validate methods for isolating tumour cells for DNA methylation analysis from rectal cancer tissue, and obtain preliminary data on genome-wide patterns of DNA methylation in rectal cancer.

This will enable future studies to identify epigenetic variation associated with response to treatment, recurrence rates and prognosis in rectal cancer.

Lay summary:

Rectal cancer is a common form of bowel cancer that usually needs an operation (surgery) to try to cure it. This can be complicated, risky and result in long-term complications. Before surgery a combination of chemotherapy and radiotherapy can be used to try and shrink the cancer but the results are variable: the cancer may shrink completely (and the person may potentially avoid surgery) or it will have no effect at all. At the moment we cannot tell how an individual will respond to this combination therapy. As the therapy itself can have adverse effects we need to find a way to identify those people who will respond very well to it and therefore safely avoid surgery.

How a person responds may be controlled by their genes, but we are not sure how. Genes are the body's instruction manual affecting the way you look and how your body works. We would like to look at the chemical switches that control the activity of the genes. Changes to these chemical switches have been seen in many different types of cancer and it may be that the pattern of these changes could help to identify how a person responds to treatment. However it is not always easy to get the right samples to look for these patterns. Our study wants to find a good reliable way of getting the rectal cancer cells we need to look at the chemical switches. This is an essential first step if we want to go on and study the possible effects of these pattern changes in rectal cancer.

Background:

Rectal cancer accounts for one third of all colorectal cancers. For patients with early disease, surgery - radical total mesorectal excision (TME) - is the mainstay of treatment[3]. Due to its anatomical location in the small confined pelvic area, surgical resection with clear margins is often difficult. Furthermore, surgery on the rectum comes with side effects, significant risk of complications, poor long-term function and reduced quality of life[3]. Additionally, most patients are found to have an intermediate probability of recurrence[3]. Current imaging modalities do not detect lymph node spread reliably, especially at a microscopic level. Consequently, in the absence of robust scientific evidence evaluating recurrence rates, the recommendation to undergo radical surgery in early stage rectal cancer is usually based on uncertainty.

NICE recommends neaodjuvant chemo-radiotherapy (NAC) prior to surgical resection in patients with advanced stage rectal tumours or nodal disease[1]. The response to the treatment is varied. A complete eradication of tumour is achieved in some patients following NAC. Other patients however, see no improvement and rely solely on surgery. Chemotherapy and radiotherapy are both associated with significant adverse effects too, including delayed wound healing, urinary and faecal incontinence and sexual dysfunction.

Therefore, being able to identify the patients that will not respond to NAC will spare them the morbidity associated with the treatment. Equally, patients that achieve complete response from NAC alone will be spared surgery and its potential complications.

There is considerable interest in the analysis of epigenetic modifications to the genome of cancer cells that do not involve a change in the nucleotide sequence. Epigenetic alterations are believed to be as important as genetic mutations in a cell's transformation to cancer, and their manipulation holds great promise for cancer prevention, detection, and therapy[2]. A variety of epigenetic mechanisms have been shown to be involved in different types of cancer, particularly the silencing of tumour suppressor genes and the activation of oncogenes by altered DNA methylation patterns[4]. Several medications which directly influence epigenetic processes are now used in various types of cancer.

Although considerable research has focused on epigenetic changes in colon cancer, especially DNA methylation differences between cancer cells and normal mucosal cells, the epigenetic and transcriptional changes specific to rectal cancer are poorly understood[3,4]. Methylomic variation in certain transcription genes and loci have been reported in rectal cancer and linked to recurrence rates and prognosis[4]. However, very few of the markers found have been replicated successfully in subsequent studies. Additionally, to our knowledge, there is no published genome-wide study looking at the differences in the epigenetic profile of rectal tumour tissue vs normal mucosa, or attempting to create an epigenetic profile for differentially methylated regions in rectal cancer.

This paucity of information may be due to limitations in the current methodologies used for the extraction and processing of rectal tissue in molecular epidemiological work. The outputs are varied and not always adequate. For DNA methylation studies, high yield rectal tumour cells should be used for analysis where possible. Colorectal adenocarcinomas will vary in histopathological features depending on the microscopic levels of inflammatory cells, neuroendocrine scattered cells and necrotic debris found within the samples. Therefore, it is impossible to quantify the content of tumour cells within a rectal tumour sample without histological workup.

Current Methodologies for tissue collection:

Currently, there is no standardised method for DNA extraction and methylation analyses for rectal cancer tumours across the limited number of studies that have been performed in this field.

A common method is DNA extraction from flash frozen cancer specimens collected from patients during surgery. In their methylation study, Leong et al used liquid nitrogen flash frozen rectal tumour samples[3]. However, the samples were not histologically assessed for heterogeneity, nor were the levels of tumour cells within the samples quantified. Therefore the discrepancy in the content of tumour cells in each sample may have subsequently affected the DNA methylation analyses.

Benard et al used formalin-fixed paraffin embedded specimens, identified as rectal cancer by a pathologist[4]. Those were subsequently micro dissected for tumour areas, although no threshold for tumour content within each specimen was determined or specified.

Jo et al used rectal tumour samples that revealed a tumour cell content of at least 60% after histological workup[5]. The methodology used for estimation of tumour cell content however was not specified. Furthermore, the threshold of 60% is deemed to be too low to produce reliable DNA methylation results.

Our study therefore aims to develop and validate methods for isolating tumour cells for DNA methylation analysis from rectal cancer tissue, and obtain preliminary data on genome-wide patterns of DNA methylation in rectal cancer.

Aims:

To develop, and validate, methods for isolating tumour cells for DNA methylation analysis from rectal cancer tissue, and obtain preliminary data on genome-wide patterns of DNA methylation in rectal cancer.

Objectives:

- Obtain matching rectal tumour and non-tumour tissues from pre-existing collected samples stored in The Exeter Tissue Bank of the NIHR Exeter Clinical Research Facility. Assess the samples histologically for integrity and cellular heterogeneity
- Dissect the samples histologically to obtain a high yield tumour DNA cells (optimally > 50% purity threshold)
- Extract DNA from the paraffin embedded tissue samples
- Perform quality control procedures to validate methodology
- Perform DNA methylation analysis on the samples using microarray technology
- Process the data using an optimised pipeline to identify patterns of differential DNA methylation in the tumour samples compared to non-tumour samples.

Methods:

Study design:
This is a methodological study utilising tissue samples currently held by the RD&E Tissue Bank (TB). Tissue remaining from this study will remain in the custody of the TB.

Inclusion criteria for this study:
Matched malignant, and benign, rectal cancer samples from the same individual.

Availability of potential tissues:
Preliminary discussions with the TB Management team have identified a sufficient number of samples would be available to facilitate this project.

Procedure:
This study aims to assess 30 matched pairs of tumour and non-tumour tissue (i.e. process 60 individual biopsies in total). Matching samples consisting of rectal tumour tissue and unaffected mucosal tissue for each patient will be selected from the TB. Sections of the flash frozen tissue received from the TB will be dissected and placed in formalin section. This will be paraffin stained and assessed histologically for integrity, heterogeneity and tumour cells content by an experienced histopathologist (**Figure 1**).

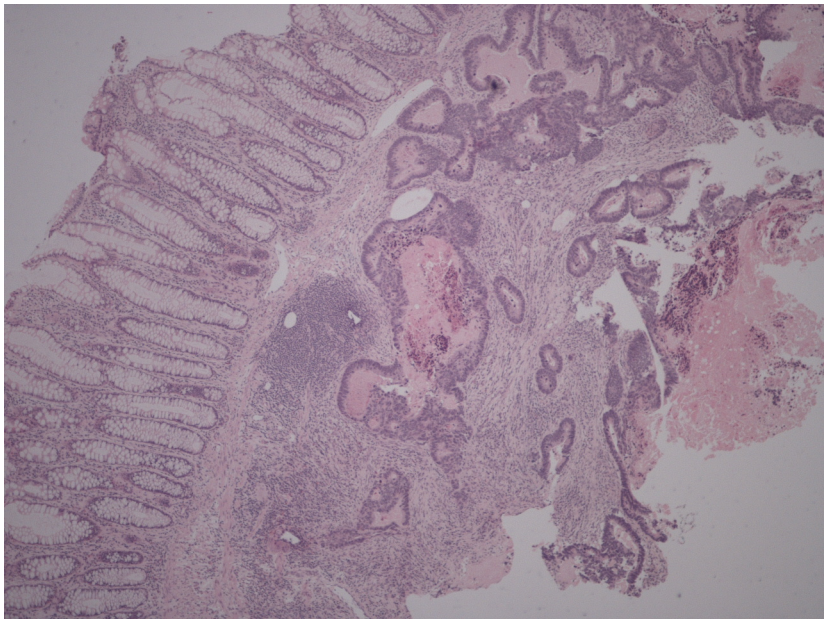


Figure 1. Paraffin stained section from a rectal tumour sample as assessed histologically.

The samples will then be dissected in order to obtain tumour tissue containing > 50% tumour cells for further processing.

Genomic DNA will be isolated from each paraffin embedded tissue sample using a method specifically designed for processing small numbers of fixed cells. DNA will be tested for quality and quantity and then treated with sodium bisulfite to convert unmethylated cytosines to uracil. DNA methylation will be quantified using the Illumina Infinium HumanMethylation 450K array, which generates a quantitative measurement of DNA methylation for >480,000 CpG sites spanning all annotated genes and other functional motifs across the genome. Data will

undergo stringent quality control and be analysed using a custom bioinformatics pipeline developed by Professor Mill's group. Following methylomic profiling, specifically differentially methylated regions (DMRs) of interest will be validated and subsequently replicated in additional samples using bisulfite-pyrosequencing. Each assay will be designed to span CpG regulatory regions likely to be affected by DNA methylation (e.g. transcption-factor binding sites).

Data storage:
All biological samples will be stored in line with the HTA[6] and MRC guidelines[7] on the handling and storage of human tissues.

All data in paper form will be kept in locked filing cabinets within the controlled access research facility.

All other study data will be kept on a password protected study database on a network drive accessible only to the study team and under the guardianship of the CI. Prof Mill's team has a dedicated secure server that will be used for data pre-processing and bioinformatics analysis after the DNA methylation arrays have been run. They routinely process samples from clinical patients and have secure data protection protocols in place.

Tissue Donor Privacy and Confidentiality:

Tissue and associated data obtained from the TB will be provided anonymised and identified only by original TB ID number.

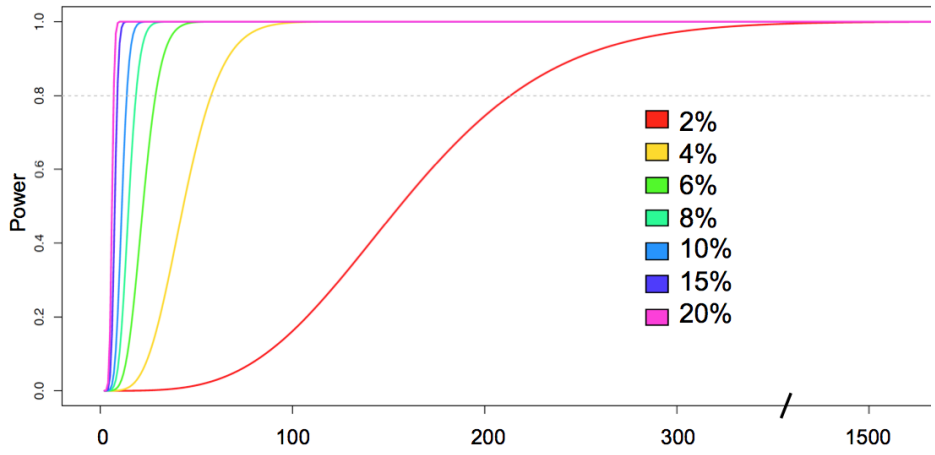
Data analysis:

DNA methylation will be assessed using the Illumina Infinium Human Methylation 450K array, which generates a quantitative measurement of DNA methylation for >480,000 CpG sites spanning all annotated genes and other functional motifs across the genome. Raw methylomic data will undergo stringent quality control assessment and be subsequently analysed using a specialized bioinformatic pipeline. Data normalisation and pre-processing will be undertaken using the *WateRmelon* package, developed by our group[8] with non-specific probes and those known to overlap polymorphic SNPs removed from subsequent analyses. Cellular heterogeneity between samples will be controlled for using cell count data and in silico algorithms designed to determine cellular composition in aggregate samples[9]. Following methylomic profiling, specific differentially methylated regions (DMRs) of interest will be confirmed and further examined by targeted bisulfite-pyrosequencing to accurately quantify DNA methylation at specific CpG sites nominated by our genome-wide analyses. Professor Mill's group has used these approaches extensively to characterise epigenetic variation in complex disease[10], and all facilities and expertise is available locally.

Sample numbers and power calculation:

We propose to use matched tumour and non-tumour tissue from 30 individuals in this study. This will provide sufficient power to address the aims of the study.

For our differential DNA methylation analysis, the study will undertake intra-individual paired analysis. Such an approach overcomes many of the confounders that can reduce power in molecular epidemiology as affected and unaffected samples are from the same donor (and thus are matched for genotype, age, sex, environmental exposures, etc.). This will represent the largest study yet undertaken in rectal cancer, and the first to take an unbiased genome-wide approach. Given the dramatic and widespread changes in the epigenome seen in cancer, we should have excellent power to identify tumour-associated differences; a power calculation using control data generated in Professor Mill's lab using the same Illumina 450K array platform that we propose to use demonstrates that our planned approach has excellent power to detect changes of the magnitude widely seen in other types of cancer – a 20% difference in DNA methylation (for which we have almost 100% power, see Figure below) would be considered a small change in this regard.



Project development and user involvement:

Clinical members of the research team have regular discussions with their patients on the importance of ensuring the right treatment for each individual. These discussions involve the need for continued research into improving treatment options. Patients have been supportive of "anything" that can improve their care and understand that to get to an end result that may improve their care on the front-line; it often needs these kinds of method development studies in the early stages. The samples obtained from the TB were donated by individuals also keen that: "it may not help me, but could help someone else further down the line"

Reporting adverse effects

In the event of an adverse event being identified during the processing of the anonymised frozen tissue, laboratory protocols will be followed and the event reported to appropriate channels.

Timescale
It is anticipated that the objectives identified above will be achieved within 3 months as outlined below:

- Obtain samples: (March 2016)
- Assess techniques for obtaining tissue cells from all samples (March 2016)
- Methylation studies (March 2016)
- Data analysis (April 2016)
- Prepare data and submit for peer-reviewed publication (April 2016)

Budget Summary and Costings:

This project will bring together NHS and UoE Medical School research staff and utilise existing NIHR infrastructure (RD&ETB) therefore research staff costs are covered.

Funding for sample analysis (approx. £175 per sample) will be covered by existing funding to Professor Mill.

Dissemination/implementation of research
Results will be written up and submitted for publication in peer-reviewed journals. Abstracts will be submitted to relevant conferences. Results will be presented to clinical/academic colleagues at regular in-house meetings.

Potential impact and benefit of the proposed research

The potential impact of this project will be both short and long term: in the short term the project will allow the identification of a specific methodology for obtaining rectal tumour cells for DNA methylation in rectal cancer biospises.

In the longer term, the availablity of a standard methodology for obtaining appropriate tissue samples will inform the development of future research projects looking at DNA methylation associated with rectal cancer, leading to an understanding of the epigenetic factors involved in rectal cancer.

Justification of research

The current research group have existing expertise in the fields of epigenetics and DNA methylation in several areas including tumour methylomic profiling, genome-wide analyse and epigenetic variation in complex disease and cancer.

End of Study
The study will finish when tissue cell collection techniques have been assessed, and sample analysis undertaken on all samples as identified in the project timescale above.

References

1- Poston GJ, Tait D, O'Connell S, Bennett A, Berendse S. Diagnosis and management of colorectal cancer: summary of NICE guidance. BMJ. 2011;343:d6751.

2- Smits, K. M., V. Melotte, H. E. Niessen, L. Dubois, C. Oberije, E. G. Troost, M. H. Starmans, P. C. Boutros, M. Vooijs, M. van Engeland and P. Lambin (2014). "Epigenetics in radiotherapy: where are we heading?" Radiother Oncol 111(2): 168-177.

3- Leong, K. J., W. Wei, L. A. Tannahill, G. M. Caldwell, C. E. Jones, D. G. Morton, G. M. Matthews and S. P. Bach (2011). "Methylation profiling of rectal cancer identifies novel markers of early-stage disease." Br J Surg 98(5): 724-734.

4- Benard, A., C. J. van de Velde, L. Lessard, H. Putter, L. Takeshima, P. J. Kuppen and D. S. Hoon (2013). "Epigenetic status of LINE-1 predicts clinical outcome in early-stage rectal cancer." Br J Cancer 109(12): 3073-3083.

5- Jo, P., K. Jung, M. Grade, L. C. Conradi, H. A. Wolff, J. Kitz, H. Becker, J. Ruschoff, A. Hartmann, T. Beissbarth, A. Muller-Dornieden, M. Ghadimi, R. Schneider-Stock and J. Gaedcke (2012). "CpG island methylator phenotype infers a poor disease-free survival in locally advanced rectal cancer." Surgery 151(4): 564-570.

6- HMGov (2004). Human Tissue Act, HMSO.

7- MRC (2001). Human Tissue and Biological Samples for use in Research.

8- Pidsley, R., Y. W. CC, M. Volta, K. Lunnon, J. Mill and L. C. Schalkwyk (2013). "A data-driven approach to preprocessing Illumina 450K methylation array data." BMC Genomics 14: 293.

9- Houseman, E. A., W. P. Accomando, D. C. Koestler, B. C. Christensen, C. J. Marsit, H. H. Nelson, J. K. Wiencke and K. T. Kelsey (2012). "DNA methylation arrays as surrogate measures of cell mixture distribution." BMC Bioinformatics 13: 86.

10- Viana, J., R. Pidsley, C. Troakes, H. Spiers, C. C. Wong, S. Al-Sarraj, I. Craig, L. Schalkwyk and J. Mill (2014). "Epigenomic and transcriptomic signatures of a Klinefelter syndrome (47,XXY) karyotype in the brain." Epigenetics 9(4): 587-599.

Appendix 4 – Supplementary Tables

Supplementary Table 1. Age Calculator results for all samples.

Sample	Age	DNA Age	Predicted Gender	Predicted Tissue	Tissue
1	49	54.16994748	male	Colon	Blood
2	49	74.92292768	male	Colon	Blood
3	71	64.9085782	female	Colon	Blood
4	71	53.01152044	female	Colon	Blood
5	75	65.0955444	female	Colon	Blood
6	75	67.92234351	female	Colon	Blood
7	76	72.19747319	female	Colon	Blood
8	76	46.56578304	male	Colon	Blood
9	79	71.75498177	female	Colon	Blood
10	79	59.93039255	female	Colon	Blood
11	77	67.4391369	female	Colon	Blood
12	77	59.72348624	female	Colon	Blood
13	74	58.42779004	female	Colon	Blood
14	74	56.67313195	female	Colon	Blood
15	71	63.85673321	female	Colon	Blood
16	71	53.7667642	female	Colon	Blood
17	67	68.43355106	male	Colon	Blood
18	67	67.9124292	male	Colon	Blood
19	63	55.4932901	female	Colon	Blood
20	63	48.42127721	female	Colon	Blood
21	65	56.90356746	female	Colon	Blood
22	65	42.72908983	female	Colon	Blood

23	79	71.54005144	male	Colon	Blood
24	79	52.87519815	male	Colon	Blood
25	56	59.96906713	male	Colon	Blood
26	56	24.84936953	male	Colon	Blood
27	72	67.97555223	male	Colon	Blood
28	72	50.03930671	male	Colon	Blood
29	78	65.22639049	male	Colon	Blood
30	78	53.58342956	male	Colon	Blood
31	45	44.53594746	male	Colon	Blood
32	45	44.7380869	male	Colon	Blood
33	49	NA	male	NA	Blood
34	49	38.49008452	male	Colon	Blood
35	66	NA	male	NA	Blood
36	66	NA	male	NA	Blood
37	62	NA	Unsure	NA	Blood
38	62	81.7601887	male	Colon	Blood
39	54	58.06056364	male	Colon	Blood
40	54	10.02421202	male	Colon	Blood
41	62	NA	female	NA	Blood
42	62	NA	female	NA	Blood
43	59	28.50444261	female	Colon	Blood
44	59	NA	female	NA	Blood
45	90	NA	male	NA	Blood
46	90	127.285128	female	Bone	Blood
47	77	61.62673632	female	Colon	Blood
48	77	66.49283439	female	Colon	Blood
49	80	78.85047564	female	Colon	Blood
50	80	27.16244925	female	Buccal	Blood
51	93	57.23264818	male	Colon	Blood

52	93	NA	male	NA	Blood
53	65	110.6556222	male	Colon	Blood
54	65	60.13889438	male	Colon	Blood
55	70	67.12283683	male	Colon	Blood
56	70	116.1693726	male	Colon	Blood
57	73	NA	male	NA	Blood
58	73	60.38563593	male	Colon	Blood
59	67	NA	male	NA	Blood
60	67	NA	Unsure	NA	Blood
61	83	NA	male	NA	Blood
62	83	NA	male	NA	Blood
63	77	58.1125704	female	Colon	Blood
64	77	38.22529811	female	Colon	Blood
65	79	NA	male	NA	Blood
66	79	NA	male	NA	Blood
67	69	69.21559324	Unsure	Colon	Blood
68	69	47.3688346	Unsure	Colon	Blood
69	81	42.89876798	male	Colon	Blood
70	81	12.68150801	Unsure	Colon	Blood
71	80	NA	Unsure	NA	Blood
72	80	NA	Unsure	NA	Blood
73	79	NA	male	NA	Blood
74	79	9.17757191	female	Colon	Blood
75	85	NA	Unsure	NA	Blood
76	85	NA	female	NA	Blood
77	44	53.0678611	female	Colon	Blood
78	44	18.19440839	female	Saliva	Blood
79	76	NA	Unsure	NA	Blood
80	76	90.66549701	Unsure	Colon	Blood

81	61	NA	male	NA	Blood
82	61	4.227320802	female	Saliva	Blood
83	56	33.82796604	male	Colon	Blood
84	56	49.11316433	male	Colon	Blood
85	54	NA	female	NA	Blood
86	54	58.97562436	female	Colon	Blood
87	67	NA	female	NA	Blood
88	67	NA	female	NA	Blood
89	45	NA	Unsure	NA	Blood
90	45	NA	female	NA	Blood

Supplementary Table 2. List of all DMPs

Probe	Chr	Chr Location	P.value	Beta effect	Map Location	Gene Name	Gene Location
cg02647878	4	47513	3.66E-10	0.449429809	154681197	RNF175;RNF175	5'UTR;1stExon
cg14650610	5	239524	7.30E-10	0.484972492	136834492	SPOCK1	5'UTR
cg13001868	17	211163	1.58E-09	0.451116562	43339223	C17orf46;LOC100133991;LOC100133991	Body;Body;TSS1500
cg13895235	7	225738	2.22E-09	0.569170792	752292	PRKAR1B;PRKAR1B;PRKAR1B;PRKAR1B;PRKAR1B;PRKAR1B	5'UTR;5'UTR;TSS200;5'UTR;5'UTR;5'UTR
cg18538668	14	295626	3.79E-09	-0.19612723	103839038		
cg03061682	15	54820	4.83E-09	0.41722365	28352098		
cg24847829	5	382403	6.29E-09	0.396895297	136834464	SPOCK1	5'UTR
cg26034516	17	398674	6.48E-09	0.346216327	76228121	LOC283999	Body
cg03576469	19	63572	6.83E-09	0.290652294	46917061	CCDC8	TSS200
cg18601167	7	296508	8.02E-09	0.559808371	752286	PRKAR1B;PRKAR1B;PRKAR1B;PRKAR1B;PRKAR1B;PRKAR1B	5'UTR;5'UTR;TSS200;5'UTR;5'UTR;5'UTR
cg05447008	6	95009	9.87E-09	0.400432306	73331114	KCNQ5;KCNQ5;KCNQ5;KCNQ5;KCNQ5	TSS1500;TSS1500;TSS1500;TSS1500;TSS1500
cg13356896	2	216007	1.15E-08	0.399185694	198650987	BOLL;BOLL	TSS200;TSS1500
cg09129067	8	154157	1.29E-08	-0.348999594	103750904		
cg23977631	2	369868	1.30E-08	0.456560177	100938799	LONRF2;LONRF2	5'UTR;1stExon
cg26238800	20	401750	1.38E-08	0.36683134	45142206	ZNF334;ZNF334	TSS200;TSS200

cg24685755	19	379961	1.39E-08	0.349845061	53758031	ZNF677	5'UTR
cg24820783	10	382004	1.41E-08	0.247143991	26504969	GAD2;GAD2	TSS1500;TSS1500
cg25480336	20	391254	1.50E-08	0.385016545	50720908	ZFP64	Body
cg16964348	7	273155	1.65E-08	0.390871674	24323799	NPY	TSS200
cg11328303	10	187365	1.94E-08	0.27109572	26505440	GAD2;GAD2;GAD2;GAD2	5'UTR;1stExon;5'UTR; 1stExon
cg09734791	8	163406	2.06E-08	0.455869807	72756155	MSC	1stExon
cg10224098	1	170996	2.26E-08	0.397396517	44873229	RNF220	5'UTR
cg04921989	2	86278	2.32E-08	0.368576617	132183100		
cg26314722	1	402967	2.36E-08	- 0.238588868	234867300		
cg04504205	20	79260	2.42E-08	- 0.323480149	45946429	ZMYND8;LOC100131496; ZMYND8;ZMYND8	Body;TSS1500; Body;Body
cg18324583	5	292503	2.48E-08	- 0.217418501	142975083		
cg11220565	20	185731	2.48E-08	0.359207899	47934802		
cg19752627	7	312333	2.51E-08	0.386976262	98467380	TMEM130;TMEM130; TMEM130	Body;Body;Body
cg03470088	1	61682	2.52E-08	0.025592041	24513939	IL28RA;IL28RA;IL28RA	TSS200;TSS200;TSS200
cg25884711	7	396613	2.52E-08	0.417868233	24323840	NPY;NPY	5'UTR;1stExon
cg06952671	2	119270	2.55E-08	0.472093568	182322268	ITGA4;ITGA4	5'UTR;1stExon
cg04377145	6	77167	2.72E-08	0.329003654	73331191	KCNQ5;KCNQ5;KCNQ5; KCNQ5;KCNQ5	TSS1500;TSS1500;TSS1500; TSS1500;TSS1500
cg14168530	2	230764	2.84E-08	0.506333301	45155991		
cg00859129	1	15911	2.87E-08	- 0.224919096	109422184	GPSM2	5'UTR
cg26958524	16	412835	2.90E-08	0.337002486	86613067	FOXL1	1stExon
cg25024074	2	385025	2.94E-08	0.311479567	182322501	ITGA4	1stExon
cg26328335	12	403139	3.26E-08	0.446031402	50354840	AQP5	TSS1500

cg26020069	6	398471	3.29E-08	-0.265475554	52382441	TRAM2	Body
cg22434409	4	348028	3.29E-08	0.406457764	21950722	KCNIP4	TSS1500
cg08266366	12	140921	3.38E-08	0.375119199	50354998	AQP5	TSS1500
cg18607529	7	296589	3.47E-08	0.408805228	50343869	IKZF1	TSS1500
cg06319475	8	108631	3.67E-08	0.38253068	145105829		
cg07589773	7	130039	3.80E-08	0.372646268	50343883	IKZF1	TSS1500
cg03020208	12	54057	3.85E-08	0.303956259	50354962	AQP5	TSS1500
cg11751707	2	193379	3.85E-08	0.273146429	38302587	CYP1B1	5'UTR
cg09220050	20	155496	3.88E-08	-0.029897082	48770642	TMEM189;TMEM189; TMEM189; TMEM189-UBE2V1	TSS1500;TSS1500; TSS1500;TSS1500
cg20415809	2	321496	3.92E-08	0.386142582	182321855	ITGA4;ITGA4	5'UTR;1stExon
cg21938148	13	341900	4.01E-08	0.479457137	110958977	COL4A1;COL4A2	Body;TSS1500
cg13405887	9	216828	4.37E-08	0.424566454	132382812	C9orf50	1stExon
cg25340966	1	389226	4.46E-08	0.340306042	119532195	TBX15	TSS200
cg17200768	13	276563	4.62E-08	0.394516994	28503373		
cg26415547	12	404524	4.82E-08	0.319863705	66583048	IRAK3;IRAK3;IRAK3;IRAK3	1stExon;1stExon;5'UTR; 5'UTR
cg07921384	2	135314	4.84E-08	0.332251121	220299740	SPEG	1stExon
cg10013343	13	167616	4.86E-08	0.502548779	29106503		
cg12628196	7	206178	4.87E-08	0.403101986	127672458	SND1;LRRC4	Body;TSS1500
cg06072021	11	104977	4.99E-08	0.516435028	128564106	FLI1;FLI1;FLI1	1stExon;5'UTR;5'UTR
cg16674351	1	269065	5.03E-08	0.34913105	121260892	LOC647121	TSS200
cg17170568	7	276109	5.05E-08	0.17708847	156433406	C7orf13;RNF32;RNF32	TSS200;1stExon;5'UTR
cg24190603	6	373133	5.05E-08	0.355341561	84418433	SNAP91;SNAP91	Body;5'UTR
cg15336765	12	249239	5.15E-08	0.42949648	50355307	AQP5;AQP5	1stExon;5'UTR
cg25223771	8	387740	5.45E-08	0.376669909	145105503		
cg00741836	20	13819	5.50E-08	0.447569313	53092233	DOK5	TSS200

cg02155398	2	38836	5.53E-08	0.364728517	45160490		
cg25680916	3	393965	5.76E-08	-0.028952666	53916015	ACTR8	1stExon
cg24924779	20	383575	5.98E-08	0.380071371	49639998	KCNG1	TSS1500
cg25773267	20	395264	6.14E-08	0.359460259	61992187	CHRNA4	Body
cg17393267	3	279350	6.15E-08	0.287395147	192127356	FGF12;FGF12	TSS1500;Body
cg12441126	7	203414	6.21E-08	0.414079829	751962	PRKAR1B;PRKAR1B;PRKAR1B;PRKAR1B;PRKAR1B;PRKAR1B	5'UTR;5'UTR;5'UTR;5'UTR;5'UTR;5'UTR
cg13267264	8	214705	6.29E-08	0.428780876	70983600	PRDM14	TSS200
cg11901272	6	195740	6.34E-08	0.379458459	29760447	HCG4	Body
cg02742906	13	49145	6.39E-08	0.340909118	112758625		
cg11947981	2	196422	6.39E-08	0.394836421	182322749	ITGA4	Body
cg26718707	10	409118	6.46E-08	-0.421586867	518370	DIP2C	Body
cg20381963	7	321042	6.48E-08	0.509017834	752238	PRKAR1B;PRKAR1B;PRKAR1B;PRKAR1B;PRKAR1B;PRKAR1B	5'UTR;5'UTR;TSS200;5'UTR;5'UTR;5'UTR
cg20107395	20	317502	6.60E-08	0.496803783	53092334	DOK5;DOK5	1stExon;5'UTR
cg24242823	7	373759	6.79E-08	0.37208258	24323675	NPY	TSS200
cg11601252	15	191089	6.92E-08	0.350932645	68122139	LBXCOR1	Body
cg04366687	8	77006	7.26E-08	0.353103589	145107199	OPLAH	Body
cg18355902	4	292935	7.28E-08	0.401077083	154681128	RNF175;RNF175	5'UTR;1stExon
cg08206318	5	139920	7.33E-08	0.299706014	134363637	PITX1	3'UTR
cg13554086	5	219602	7.39E-08	0.389158636	76507100	PDE8B;PDE8B;PDE8B;PDE8B;PDE8B	Body;Body;Body;Body; Body
cg25645268	4	393529	7.81E-08	0.305821609	154710598	SFRP2	TSS1500
cg17101450	10	275064	7.85E-08	0.377444704	102900365		

cg00321614	5	5977	7.88E-08	-0.174805125	172856932		
cg14215472	17	231713	7.91E-08	0.350417461	27940404	ANKRD13B	Body
cg18335068	19	292663	7.99E-08	0.315843801	53757910	ZNF677	5'UTR
cg09296001	7	156745	8.21E-08	0.521700368	127672564	SND1	Body
cg03424342	3	60835	8.35E-08	0.279116055	120169783	FSTL1;FSTL1	5'UTR;1stExon
cg26593267	13	407219	8.47E-08	0.341347306	113764871	F7;F7	Body;Body
cg10120816	6	169297	8.61E-08	0.256360286	99296305		
cg09871471	1	165522	8.62E-08	0.41166646	121260900	LOC647121	TSS200
cg03147907	3	56232	9.20E-08	-0.336528803	62926690		
cg10928466	19	181453	9.24E-08	0.108647309	11353961	DOCK6	Body
cg07719492	8	132095	9.24E-08	0.315052192	70983348	PRDM14	5'UTR
cg21995919	2	342525	9.27E-08	0.414609207	182322279	ITGA4;ITGA4	5'UTR;1stExon
cg23383871	20	361615	9.31E-08	0.372836732	47934987		
cg10770742	7	179135	9.31E-08	0.372960807	151107285	WDR86	TSS200
cg04023150	1	71075	9.33E-08	0.447449018	44873064	RNF220	5'UTR
cg17228900	6	277000	9.42E-08	0.453670006	391764	IRF4;IRF4	5'UTR;1stExon
cg27532621	1	421231	9.61E-08	0.230813829	164593763	PBX1	Body
cg10065823	9	168383	1.00E-07	0.29923162	96108467	C9orf129;C9orf129	1stExon;5'UTR
cg11573679	2	190672	1.03E-07	0.475379695	68546467	CNRIP1;CNRIP1	1stExon;1stExon
cg23572908	7	364168	1.05E-07	0.370648121	158937969	VIPR2	TSS1500
cg01440841	4	26322	1.06E-07	0.351171663	154681066	RNF175;RNF175	5'UTR;1stExon
cg05946309	16	102971	1.07E-07	-0.208488839	85926085		
cg05288172	8	92328	1.07E-07	-0.26607755	103751006		
cg18918321	8	300921	1.08E-07	0.2622105	41424524		
cg02177231	1	39206	1.09E-07	0.333194537	119529930	TBX15	5'UTR
cg16437728	11	265286	1.09E-07	0.364624499	7273046	SYT9	TSS200

cg08957069	6	151718	1.11E-07	0.316843199	28743700		
cg23934404	13	369192	1.13E-07	0.418802979	112758491		
cg23933289	1	369171	1.14E-07	0.233664502	178998656	FAM20B	5'UTR
cg11338643	6	187521	1.16E-07	0.409405033	166580983	T	Body
cg08750504	2	148450	1.20E-07	0.387758312	172946193		
cg09073398	5	153372	1.20E-07	0.385215757	168727762	SLIT3;SLIT3	5'UTR;1stExon
cg14337134	7	233929	1.21E-07	0.176560163	102920323	DPY19L2P2;DPY19L2P2	Body;Body
cg21232488	6	332729	1.23E-07	-0.279429477	30079203	TRIM31	Body
cg16504626	8	266227	1.25E-07	0.384265106	57070013		
cg01618245	20	29403	1.29E-07	-0.237429382	61990279	CHRNA4	Body
cg14443519	6	235755	1.30E-07	0.433782606	29760410	HCG4	Body
cg27141850	2	415570	1.30E-07	0.379147268	20869434	GDF7	Body
cg14015706	9	228020	1.30E-07	0.502648076	132382433	C9orf50	1stExon
cg24171907	2	372892	1.32E-07	0.447022617	68546579	CNRIP1;CNRIP1; CNRIP1;CNRIP1	1stExon;1stExon;5'UTR; 5'UTR
cg20652954	20	324617	1.33E-07	-0.197332275	61716293		
cg10210594	1	170794	1.34E-07	0.330872402	208132787		
cg12433277	7	203284	1.34E-07	0.314504362	151106990	WDR86;WDR86	5'UTR;1stExon
cg02700626	11	48369	1.35E-07	0.240343059	64739320		
cg20295442	8	319933	1.37E-07	0.477998364	67344665	ADHFE1	TSS200
cg08332074	16	142027	1.37E-07	0.309815416	51189941		
cg21647227	1	338120	1.37E-07	0.363506599	119527111	TBX15	5'UTR
cg06401021	6	109895	1.37E-07	0.416252367	55443868	HMGCLL1;HMGCLL1; HMGCLL1;HMGCLL1	1stExon;5'UTR;1stExon; 5'UTR
cg23641267	11	365125	1.39E-07	0.251295468	58343791	LPXN;LPXN	Body;TSS1500
cg23092823	1	357374	1.39E-07	0.245227874	53528612	PODN	Body

cg25771271	1	395231	1.41E-07	0.319163644	119550191		
cg27341128	20	418487	1.41E-07	0.333202735	53092259	DOK5	TSS200
cg14595003	3	238570	1.41E-07	0.383497013	129694156	TRH	5'UTR
cg25189564	7	387224	1.43E-07	0.29061549	158938051	VIPR2	TSS1500
cg24645214	8	379359	1.45E-07	0.30076221	54789978	RGS20	Body
cg07878486	19	134591	1.46E-07	0.326438411	58951885	ZNF132	TSS1500
cg18303242	17	292176	1.48E-07	-0.368987913	25879250	KSR1	5'UTR
cg12619536	1	206042	1.49E-07	0.41578936	108508067	VAV3	TSS1500
cg17892556	19	286136	1.49E-07	0.466224169	12267464	ZNF625;ZNF625	1stExon;5'UTR
cg04415599	19	77844	1.49E-07	0.239895854	37464508		
cg08516516	5	144823	1.50E-07	0.300889602	115152492	CDO1	TSS200
cg02065637	20	37250	1.51E-07	0.363052106	61809035	MIR124-3	TSS1500
cg08569799	5	145637	1.52E-07	0.294087308	1886828		
cg17226446	4	276961	1.53E-07	0.139640018	154408845	KIAA0922;KIAA0922	Body;Body
cg25060829	6	385522	1.56E-07	0.309825561	28367571	ZSCAN12;ZSCAN12	TSS200;TSS200
cg19991022	20	315982	1.56E-07	0.545169429	53091929	DOK5	TSS1500
cg13850380	1	224969	1.60E-07	0.317619683	1475143	C1orf70	Body
cg15745900	8	255165	1.62E-07	0.381411419	68864549	PREX2;PREX2	TSS200;TSS200
cg27200446	6	416403	1.62E-07	0.57207624	41606439	MDFI	5'UTR
cg07790085	13	133165	1.64E-07	-0.253479546	29597447	MTUS2	TSS1500
cg20912169	8	328051	1.64E-07	0.479902615	67344720	ADHFE1;ADHFE1	5'UTR;1stExon
cg00250422	15	4717	1.65E-07	0.415840996	28352347		
cg09632907	4	161903	1.68E-07	0.242282767	54969963		
cg18435449	19	294072	1.70E-07	0.331667026	58095445	ZIK1	TSS200
cg17371081	11	279020	1.70E-07	0.277851676	20690957	NELL1;NELL1	TSS200;TSS200
cg03064067	12	54860	1.70E-07	0.394769427	85306916	SLC6A15;SLC6A15; SLC6A15	TSS1500;TSS1500;TSS1500

cg12868067	12	209399	1.71E-07	0.379758435	128752246	TMEM132C	Body
cg13389502	17	216514	1.72E-07	0.161722774	1961440	HIC1;HIC1	Body;Body
cg06427779	5	110268	1.73E-07	0.319148477	54180079		
cg27188703	12	416248	1.76E-07	0.329107422	50297581	FAIM2;FAIM2	1stExon;5'UTR
cg11903130	10	195775	1.82E-07	0.34806647	26506751	GAD2;GAD2	Body;Body
cg18921980	2	300977	1.82E-07	0.27771105	175594943		
cg16306898	1	263170	1.83E-07	0.477786305	1475675	C1orf70	1stExon
cg02640612	8	47384	1.84E-07	0.282845368	53853444	NPBWR1	1stExon
cg16366473	3	264117	1.88E-07	0.412447519	192126849	FGF12;FGF12	TSS200;Body
cg22862480	10	353888	1.91E-07	0.486456346	7450355	SFMBT2	5'UTR
cg14658804	5	239677	1.92E-07	0.330784284	168728213	SLIT3	TSS200
cg23575688	11	364206	1.93E-07	-0.206365044	119486443		
cg21013866	14	329498	1.95E-07	0.357490982	23834985	EFS;EFS	TSS200;TSS200
cg24663256	4	379578	1.96E-07	0.382297975	21950307	KCNIP4;KCNIP4	5'UTR;1stExon
cg22663389	17	350997	1.97E-07	-0.352405177	57929274		
cg01046104	19	19396	1.98E-07	0.26295962	58095588	ZIK1	TSS200
cg04803843	15	84191	1.99E-07	0.423122961	28351906		

Supplementary Table 3 The 53 DMRs annotated to genes also identified in the list of DMPs.

Region	Gene	N.Probes	P Value	Epigenetic/Cancer Literature
chr6:391114-392131	IRF4	11	1.72E-14	Related to the interferon signalling pathway IFN, which has been found to play a role in rectal cancer (Martha Slattery)
chr8:67344553-67345006	ADHFE1	9	1.13E-15	Promoter hypermethylation frequently present in CRC.
chr11:128693677-128694679	FLI1	9	7.83E-07	Promising biomarker for rectal cancer as part of a specific DNA methylation signature (Vymetalkova,V, 2016)
chr1:25257505-25258082	RUNX3	13	3.86E-05	Part of CIMP Panel for CRC
chr1:108507078-108508207	VAV3; VAV3-AS1	8	9.67E-13	Overexpressed in CRC and is correlated with tumour metastases and invasion (Yih-Huei)
chr19:58095011-58095659	ZIK1	9	2.00E-14	Methylated in azoxymethane colon cancer model (Borinstein SC, 2010)
chr7:751830-752456	PRKAR1B	9	1.24E-19	Included in Top 10 DMPs
chr5:136833893-136834492	SPOCK1	7	1.05E-13	Included in Top 10 DMPs
chr4:154680808-154681619	RNF175	8	5.39E-12	Included in Top 10 DMPs
chr2:68546467-68547256	CNRIP1	9	4.37E-12	Included in a novel combined biomarker panel for detection of CRC and premalignant malinomas (Guro E Lind)(Wang Pei)

chr2:100938799-100939477	LONRF2	6	4.88E-09	Hypermethylation thought to contribute to tumourigenesis or rectal cancer (Hua Y, 2017)
chr7:24323128-24323939	NPY	10	1.82E-17	hypermethylated in CRC and promising biomarker (Roperch Jean PierrE)
chr6:41606317-41607326	MDFI	8	5.60E-15	Hypermethylated in CRC (Lin PC,2014)
chr19:12267308-12267796	ZNF625; ZNF20	6	2.36E-09	Hypermethylated in CRC (Lin PC, 2015)
chr10:26504950-26505503	GAD2	9	1.97E-13	Hypermethylated in CRC (Hai Li, 2012)
chr17:1961109-1961778	HIC1	5	1.58E-07	Hypermethylated in CRC (H P Mohamed)
chr6:73331114-73331680	KCNQ5	7	3.56E-13	Hypermethylated in CRC (Ashktorab, 2016).
chr5:168728076-168728586	SLIT3	8	1.67E-09	Frequently methylated in CRC and significantly associated with TNM stage, lymph node metastasis and differentiation (Tingting huang)
chr2:182321489-182322841	ITGA4	11	3.88E-18	Frequently methylated in CRC and adenoma tissues (Xie Zhang).
chr4:154710224-154710961	SFRP2	17	2.45E-12	Frequently methylated in CRC and adenoma tissues (Xie Zhang).
chr4:154713235-154713789	SFRP2	6	6.63E-07	Frequently methylated in CRC and adenoma tissues (Xie Zhang).
chr6:84419189-84419360	SNAP91	9	9.39E-09	Downregulated due to promoter hypermethylation in around 60% of colorectal cancer and may be associated with treatment resistance. (Andrew Beggs)
chr5:115152019-115152938	CDO1	12	3.58E-10	Considered to play a role as potential biomarker for CRC detection. (Yamashita K) (Vedeld HM)

chr14:23834710-23835212	EFS	9	6.24E-13	Associated with prediction of biochemical, local, and systemic recurrence of prostate cancer (Vanaja DK, 2009)
chr3:129693370-129694683	TRH	11	7.14E-19	_
chr1:119526783-119528848	TBX15	14	8.02E-19	_
chr5:134362967-134363973	PITX1	9	2.38E-16	_
chr8:70982867-70983760	PRDM14	10	3.32E-16	_
chr11:7272653-7273735	SYT9	11	8.23E-16	_
chr8:68864012-68865187	PREX2	14	4.04E-15	_
chr1:119531625-119532925	TBX15	17	9.24E-15	_
chr20:53091739-53092984	DOK5	10	2.18E-13	_
chr4:21949853-21951046	KCNIP4	10	3.38E-13	_
chr11:20690628-20691429	NELL1	12	6.23E-13	_
chr11:2160875-2162510	IGF2; INS-IGF2	41	8.06E-13	_
chr8:72755568-72756341	MSC; MSC-AS1	6	8.26E-13	_

chr12:50297477-50297945	FAIM2	6	1.41E-12	_
chr8:53852030-53852661	NPBWR1	8	1.59E-12	_
chr3:192126584-192126996	FGF12	7	1.66E-12	_
chr20:61809348-61809932	MIR124-3	10	1.70E-12	_
chr6:28367070-28367898	ZSCAN12	10	2.52E-12	_
chr20:45141911-45142336	ZNF334	7	7.55E-12	_
chr19:53757677-53758609	ZNF677	9	1.04E-11	_
chr7:151106990-151107807	WDR86-AS1; WDR86	7	1.18E-11	_
chr6:166581929-166582585	T	10	5.90E-11	_
chr2:162283189-162284206	TBR1	9	7.61E-11	_
chr5:134366162-134367394	PITX1	8	1.18E-10	_
chr10:26506353-26506933	GAD2	5	1.83E-10	_
chr10:7453329-7454063	SFMBT2	9	2.77E-10	_

chr19:58951599-58952249	ZNF132	9	3.87E-10	_
chr6:166580460-166581538	T	7	4.13E-10	_
chr1:119521928-119522855	TBX15	5	7.62E-10	_
chr5:76506484-76507680	PDE8B	8	1.44E-09	_
chr3:192232468-192233231	FGF12	5	9.14E-09	_
chr1:164290179-164290833	PBX1	8	1.81E-08	_
chr12:128752040-128752586	TMEM132C	5	5.97E-08	_
chr10:7452242-7452950	SFMBT2	5	1.02E-07	_
chr6:30080189-30080782	TRIM31-AS1; TRIM31	7	2.02E-07	_
chr6:55443757-55444488	HMGCLL1	10	2.41E-07	_
chr2:220299484-220300242	SPEG	11	1.83E-06	_
chr8:70946891-70947440	PRDM14	5	1.20E-05	_
chr8:72756656-72757004	MSC; MSC-AS1	8	1.34E-05	_

chr1:119530600-119531122	TBX15	6	2.82E-05	_
chr12:128850196-128850696	TMEM132C	6	3.02E-04	_

Supplementary Table 4. List of All DMRs.

CHR	Start	End	N of Probes	Sildak P Value	Gene Name	CpG Island Extension Name
chr7	751830	752456	9	1.24E-19	PRKAR1B	CpG: 54
chr3	129693370	129694683	11	7.14E-19	TRH	CpG: 155
chr1	119526783	119528848	14	8.02E-19	TBX15	CpG: 35
chr10	118031632	118033902	24	8.33E-19	GFRA1	CpG: 275
chr2	182321489	182322841	11	3.88E-18	ITGA4	CpG: 119
chr7	24323128	24323939	10	1.82E-17	NPY	CpG: 123
chr8	69243285	69244734	12	1.66E-16	C8orf34-AS1;C8orf34	CpG: 83
chr6	133561224	133562776	41	1.76E-16	EYA4	CpG: 138
chr5	134362967	134363973	9	2.38E-16	PITX1	CpG: 174
chr8	70982867	70983760	10	3.32E-16	PRDM14	CpG: 253
chr2	45159504	45160554	10	3.88E-16	SIX3-AS1	CpG: 18
chr18	12253749	12254976	20	6.91E-16	CIDEA	CpG: 85
chr11	7272653	7273735	11	8.23E-16	SYT9	CpG: 139
chr15	28351619	28352558	6	8.77E-16	HERC2	CpG: 146
chr8	67344553	67345006	9	1.13E-15	ADHFE1	CpG: 51
chr13	78492306	78494462	48	1.13E-15	EDNRB	CpG: 70
chr7	93519220	93520566	20	1.52E-15	TFPI2; LOC105375401	CpG: 70
chr8	68864012	68865187	14	4.04E-15	PREX2	CpG: 48
chr6	41606317	41607326	8	5.60E-15	MDFI	CpG: 168; CpG: 20

chr16	51183988	51185772	25	7.59E-15	SALL1	CpG: 366
chr1	119531625	119532925	17	9.24E-15	TBX15	CpG: 19
chr6	391114	392131	11	1.72E-14	IRF4	CpG: 261
chr19	58095011	58095659	9	2.00E-14	ZIK1	CpG: 88
chr17	32906991	32908285	12	7.12E-14	TMEM132E	CpG: 276
chr5	87973439	87974547	12	8.39E-14	LINC00461	CpG: 34
chr2	144694405	144695257	6	8.72E-14	LOC101928386	CpG: 69
chr5	136833893	136834492	7	1.05E-13	SPOCK1	CpG: 134
chr3	140769691	140771372	19	1.48E-13	SPSB4	CpG: 103
chr2	45155201	45156207	9	1.60E-13	SIX3-AS1	CpG: 177
chr10	26504950	26505503	9	1.97E-13	GAD2	CpG: 242
chr10	129535138	129535898	9	2.00E-13	FOXI2	CpG: 238
chr20	53091739	53092984	10	2.18E-13	DOK5	CpG: 96
chr12	5018229	5019482	15	2.67E-13	KCNA1	CpG: 202
chr4	4863678	4864902	11	2.86E-13	MSX1	CpG: 40
chr4	21949853	21951046	10	3.38E-13	KCNIP4	CpG: 24; CpG: 36
chr6	73331114	73331680	7	3.56E-13	KCNQ5	CpG: 175
chr7	142494148	142495098	15	4.29E-13	PRSS3P2	CpG: 68
chr14	100437951	100438957	7	4.99E-13	EML1	CpG: 128
chr1	179711583	179712740	13	5.60E-13	FAM163A	CpG: 199
chr12	104697193	104697631	12	5.63E-13	TXNRD1; TXNRD1; EID3	CpG: 56
chr1	111217194	111218287	12	5.91E-13	KCNA3	CpG: 190
chr11	20690628	20691429	12	6.23E-13	NELL1	CpG: 114
chr14	23834710	23835212	9	6.24E-13	EFS	CpG: 131

chr11	2160875	2162510	41	8.06E-13	IGF2; INS-IGF2; IGF2-AS;IGF2;IGF2	CpG: 302
chr8	72755568	72756341	6	8.26E-13	MSC;MSC-AS1	CpG: 88
chr2	56150255	56151179	12	9.33E-13	EFEMP1	CpG: 70
chr1	108507078	108508207	8	9.67E-13	VAV3; VAV3-AS1	CpG: 78
chr2	98962900	98963768	8	1.17E-12	CNGA3	CpG: 114
chr12	50297477	50297945	6	1.41E-12	FAIM2	CpG: 40
chr5	178017260	178018186	13	1.43E-12	COL23A1	CpG: 111
chr20	44746392	44747006	11	1.44E-12	CD40	CpG: 18
chr12	85305219	85307152	18	1.49E-12	SLC6A15	CpG: 17; CpG: 16; CpG: 44
chr13	28500882	28501413	5	1.51E-12	PDX1	CpG: 23
chr11	128563483	128564180	9	1.53E-12	FLI1; SENCR; FLI1;FLI1	CpG: 194
chr4	96470053	96471143	12	1.54E-12	UNC5C	CpG: 94
chr8	53852030	53852661	8	1.59E-12	NPBWR1	CpG: 245
chr10	15760947	15762312	13	1.62E-12	ITGA8	CpG: 77
chr3	192126584	192126996	7	1.66E-12	FGF12	CpG: 176
chr20	61809348	61809932	10	1.70E-12	MIR124-3	CpG: 424
chr4	154710224	154710961	17	2.45E-12	SFRP2	CpG: 112
chr6	28367070	28367898	10	2.52E-12	ZSCAN12	CpG: 40
chr20	41818356	41819125	8	2.74E-12	PTPRT	CpG: 177
chr21	27011139	27012615	12	3.20E-12	JAM2	CpG: 64
chr15	83775850	83776446	8	3.40E-12	TM6SF1	CpG: 115
chr6	31685392	31685575	7	3.63E-12	LY6G6D	CpG: 26
chr4	110223040	110224288	19	3.70E-12	COL25A1	CpG: 112

chr20	25062254	25063052	10	4.04E-12	VSX1	CpG: 116
chr2	68546467	68547256	9	4.37E-12	CNRIP1	CpG: 75
chr1	119548527	119549312	6	4.51E-12	LOC105378933	CpG: 176
chr7	38670412	38671001	8	4.75E-12	AMPH	CpG: 70
chr6	30078754	30079662	19	4.93E-12	TRIM31-AS1; TRIM31; TRIM31	CpG: 15
chr4	154680808	154681619	8	5.39E-12	RNF175	CpG: 66
chr4	166794786	166796183	13	6.92E-12	TLL1	CpG: 68
chr20	45141911	45142336	7	7.55E-12	ZNF334	CpG: 36
chr21	38076709	38077971	9	7.67E-12	SIM2	CpG: 70
chr12	54392988	54394212	18	7.90E-12	HOXC9; HOXC-AS1	CpG: 118
chr13	95364510	95364993	11	9.36E-12	SOX21-AS1; SOX21	CpG: 276
chr19	53757677	53758609	9	1.04E-11	ZNF677	CpG: 37
chr8	97172825	97173722	7	1.08E-11	GDF6	CpG: 113
chr20	37302636	37303484	7	1.14E-11	ARHGAP40	CpG: 59
chr7	151106990	151107807	7	1.18E-11	WDR86-AS1; WDR86	CpG: 157
chr5	140864020	140864834	8	1.19E-11	PCDHGA8; PCDHGA1; PCDHGB7; PCDHGB5; PCDHGC4; PCDHGA11; PCDHGA12; PCDHGA9; PCDHGA7; PCDHGA10; PCDHGA5; PCDHGA2; PCDHGA3; PCDHGC3; PCDHGB6; PCDHGB4; PCDHGB3; PCDHGB2; PCDHGB1; PCDHGA6; PCDHGA4	CpG: 22
chr3	196065106	196065688	12	1.66E-11	TM4SF19-TCTEX1D2; TM4SF19	CpG: 87
chr16	77822083	77823221	12	1.70E-11	VAT1L	CpG: 72
chr6	87646740	87647644	15	2.14E-11	HTR1E	CpG: 43
chr6	30737007	30737455	6	2.16E-11	IER3	CpG: 161

chr10	57389676	57391271	10	2.43E-11	MTRNR2L5	CpG: 75
chr17	43339040	43339831	12	2.45E-11	SPATA32; MAP3K14-AS1; MAP3K14-AS1	CpG: 65
chr5	83678796	83680326	11	2.53E-11	EDIL3	CpG: 111
chr2	223162666	223163809	14	2.54E-11	CCDC140;PAX3	CpG: 75
chr11	18813191	18813556	5	2.65E-11	PTPN5	CpG: 115
chr16	55689865	55690564	8	2.77E-11	SLC6A2	CpG: 120
chr17	6616351	6617580	12	3.15E-11	SLC13A5	CpG: 92
chr11	64739253	64739852	8	3.23E-11	C11orf85	CpG: 70
chr4	17782643	17783502	12	3.38E-11	FAM184B	CpG: 94
chr6	10881782	10882336	11	3.49E-11	GCM2	CpG: 23
chr8	55370192	55370579	10	5.00E-11	SOX17	CpG: 233
chr6	166581929	166582585	10	5.90E-11	T	CpG: 270
chr8	143532613	143533532	7	6.32E-11	ADGRB1	CpG: 331
chr14	97684772	97686001	11	7.08E-11	LOC101929241	CpG: 74
chr17	6679254	6679781	12	7.34E-11	FBXO39	CpG: 49
chr13	108518955	108521062	16	7.35E-11	FAM155A	CpG: 139
chr2	162283189	162284206	9	7.61E-11	TBR1	CpG: 85
chr5	178367621	178368620	11	7.75E-11	ZNF454	CpG: 91
chr4	184826215	184827086	12	7.77E-11	STOX2	CpG: 197
chr1	221067662	221069136	9	7.88E-11	HLX	CpG: 66; CpG: 31
chr14	85995655	85996873	15	7.94E-11	FLRT2	CpG: 45
chr3	11033866	11035357	12	8.43E-11	SLC6A1	CpG: 87
chr17	33775295	33776554	13	8.45E-11	SLFN13	CpG: 40; CpG: 28

chr7	130130478	130131887	33	8.92E-11	MEST; MESTIT1; MEST	CpG: 177
chr10	91294781	91295855	14	9.01E-11	SLC16A12	CpG: 100
chr1	119549626	119550484	5	9.02E-11	LOC105378933	CpG: 176
chr3	134369339	134370241	13	9.65E-11	KY	CpG: 79
chr21	34397654	34398532	11	9.92E-11	OLIG2	CpG: 348
chr6	125283726	125284659	11	1.04E-10	RNF217-AS1; RNF217	CpG: 128
chr8	41166530	41167278	8	1.04E-10	SFRP1	CpG: 147
chr2	119602212	119603969	13	1.10E-10	EN1	CpG: 128
chr5	134366162	134367394	8	1.18E-10	PITX1	CpG: 51
chr13	79169714	79171469	15	1.27E-10	RNF219-AS1	CpG: 95
chr21	38080526	38081976	8	1.33E-10	SIM2	CpG: 153
chr4	5052795	5053596	8	1.33E-10	STK32B	CpG: 97
chr2	120281354	120281999	6	1.35E-10	SCTR	CpG: 67
chr6	29594830	29595661	15	1.40E-10	GABBR1	CpG: 39
chr4	111533267	111533951	7	1.41E-10	PITX2	CpG: 59
chr13	88323940	88324879	9	1.44E-10	SLITRK5	CpG: 92
chr11	105480771	105481863	15	1.51E-10	GRIA4	CpG: 31
chr8	89339404	89340750	11	1.51E-10	MMP16	CpG: 37
chr7	127672152	127672658	7	1.53E-10	SND1	CpG: 156
chr4	1397664	1398798	6	1.62E-10	NKX1-1	CpG: 534
chr7	43151828	43152542	8	1.76E-10	HECW1	CpG: 124
chr7	19156621	19158747	30	1.82E-10	TWIST1	CpG: 152
chr10	26506353	26506933	5	1.83E-10	GAD2	CpG: 242
chr2	467799	468413	5	1.89E-10	FAM150B	CpG: 79

chr20	61638149	61638588	8	1.98E-10	BHLHE23	CpG: 272
chr21	36042170	36042752	5	1.99E-10	CLIC6	CpG: 204
chr11	44330903	44333192	46	2.11E-10	ALX4	CpG: 111; CpG: 36
chr6	29759947	29761044	19	2.20E-10	LOC554223; HCG4	CpG: 78
chr18	44337664	44338147	6	2.20E-10	ST8SIA5	CpG: 47
chr19	37464066	37464633	5	2.22E-10	ZNF568	CpG: 46
chr1	20878540	20879813	11	2.28E-10	FAM43B	CpG: 188
chr2	74742309	74743243	6	2.28E-10	TLX2	CpG: 288
chr2	154727907	154728468	7	2.41E-10	GALNT13	CpG: 37
chr1	91184290	91185749	11	2.45E-10	BARHL2	CpG: 36; CpG: 124
chr4	778924	779880	7	2.53E-10	CPLX1	CpG: 172
chr6	117086479	117087083	9	2.56E-10	FAM162B	CpG: 112
chr15	79382548	79383560	10	2.68E-10	RASGRF1	CpG: 195
chr10	7453329	7454063	9	2.77E-10	SFMBT2	CpG: 418
chr1	77333749	77334256	7	2.86E-10	ST6GALNAC5	CpG: 112
chr2	38302230	38302892	5	3.07E-10	CYP1B1	CpG: 300
chr5	115152019	115152938	12	3.58E-10	CDO1	CpG: 122
chr19	58951599	58952249	9	3.87E-10	ZNF132	CpG: 81
chr6	166580460	166581538	7	4.13E-10	T	CpG: 270
chr19	56879207	56879994	13	4.30E-10	ZNF542P	CpG: 57
chr4	4859772	4860698	20	4.39E-10	MSX1	CpG: 41
chr6	99295470	99296305	5	4.52E-10	POU3F2	CpG: 51
chr14	102247610	102248253	5	4.66E-10	PPP2R5C	CpG: 64

chr2	115919785	115920612	6	4.68E-10	DPP10	CpG: 203
chr4	157996959	157997750	13	5.38E-10	GLRB	CpG: 62
chr2	139537544	139537845	5	5.46E-10	NXPH2	CpG: 100
chr14	52535758	52536436	11	5.56E-10	NID2	CpG: 160
chr18	5894665	5895539	6	5.64E-10	TMEM200C	CpG: 120
chr20	37352612	37353158	9	5.67E-10	SLC32A1	CpG: 380
chr7	121940216	121940992	7	5.73E-10	FEZF1	CpG: 53
chr2	106681831	106682640	8	5.74E-10	C2orf40	CpG: 48
chr7	155166205	155167511	10	5.78E-10	BLACE	CpG: 277
chr10	106399877	106400880	11	5.88E-10	SORCS3	CpG: 295
chr1	942161	942696	9	5.97E-10	ISG15	CpG: 413
chr19	44952417	44952808	9	6.19E-10	ZNF229	CpG: 37
chr1	110610899	110612044	8	6.59E-10	ALX3	CpG: 264
chr8	65710790	65711916	12	6.96E-10	CYP7B1	CpG: 70
chr10	134900906	134902278	13	7.54E-10	ADGRA1	CpG: 155
chr16	10276081	10277317	13	7.59E-10	GRIN2A	CpG: 278
chr1	119521928	119522855	5	7.62E-10	TBX15	CpG: 36
chr8	75896720	75897310	8	8.20E-10	CRISPLD1	CpG: 60
chr10	102279330	102279954	10	8.37E-10	SEC31B	CpG: 66
chr2	29337946	29338636	11	8.41E-10	CLIP4	CpG: 103
chr6	62995876	62996299	9	9.15E-10	KHDRBS2	CpG: 45
chr2	119599189	119600002	7	9.52E-10	EN1	CpG: 103; CpG: 20
chr14	29235904	29236535	13	1.13E-09	FOXG1	CpG: 77
chr3	142839578	142840240	6	1.37E-09	CHST2	CpG: 303

chr13	51417469	51418053	9	1.42E-09	DLEU7-AS1; DLEU7	CpG: 69
chr5	76506484	76507680	8	1.44E-09	PDE8B	CpG: 117
chr3	68981852	68982098	9	1.49E-09	FAM19A4	CpG: 105
chr7	1272255	1272710	9	1.49E-09	UNCX	CpG: 923
chr1	16084445	16085642	9	1.50E-09	FBLIM1	CpG: 76
chr6	100911240	100912168	13	1.64E-09	SIM1	CpG: 109
chr6	1390361	1391265	6	1.66E-09	FOXF2; MIR6720	CpG: 184
chr5	168728076	168728586	8	1.67E-09	SLIT3	CpG: 96
chr4	156587884	156589291	12	1.82E-09	GUCY1A3	CpG: 89
chr6	133563342	133564066	7	1.99E-09	EYA4	CpG: 138; CpG: 16
chr19	12267308	12267796	6	2.36E-09	ZNF625; ZNF625;ZNF625-ZNF20	CpG: 58
chr13	23734299	23734763	5	2.38E-09	SGCG	CpG: 93
chr8	77594595	77595359	7	2.49E-09	ZFHX4-AS1; ZFHX4	CpG: 50
chr2	95690821	95692010	8	2.54E-09	MAL	CpG: 116
chr4	122301573	122302331	11	2.59E-09	QRFPR	CpG: 67
chr6	5994235	5995371	9	2.63E-09	NRN1	CpG: 20
chr2	209271164	209272057	10	2.81E-09	PTH2R	CpG: 70
chr6	108488335	108489292	7	2.92E-09	NR2E1	CpG: 318
chr5	78985425	78986160	12	2.96E-09	CMYA5	CpG: 35
chr6	32187918	32189558	12	3.03E-09	NOTCH4	CpG: 92
chr4	5709858	5710411	9	3.18E-09	EVC2	CpG: 56
chr21	22369802	22370864	9	3.19E-09	NCAM2	CpG: 92
chr6	85473773	85474595	11	3.27E-09	TBX18	CpG: 129

chr13	102068989	102069556	9	3.28E-09	NALCN	CpG: 99
chr1	200007403	200008026	5	3.39E-09	NR5A2	CpG: 72
chr5	5139334	5140029	16	3.40E-09	CTD-2297D10.2	CpG: 140
chr12	103351855	103352694	9	4.02E-09	ASCL1	CpG: 105
chr1	91301204	91301962	6	4.36E-09	LOC105378853	CpG: 86
chr22	19137874	19138475	6	4.52E-09	GSC2	CpG: 202
chr20	3218145	3218905	11	4.53E-09	SLC4A11	CpG: 247
chr1	22140769	22141400	7	4.65E-09	LDLRAD2	CpG: 49
chr11	111410935	111411381	9	4.66E-09	LAYN	CpG: 110
chr2	100938799	100939477	6	4.88E-09	LONRF2	CpG: 153
chr11	94134015	94135029	12	5.26E-09	GPR83	CpG: 64
chr13	96204493	96204873	8	5.30E-09	CLDN10	CpG: 68
chr20	23029640	23030343	5	5.35E-09	THBD	CpG: 312
chr1	98510865	98511792	9	5.59E-09	MIR137HG; MIR137; MIR2682	CpG: 73
chr16	86598792	86599976	9	5.71E-09	FOXC2-AS1	CpG: 283
chr6	29520698	29521803	38	5.99E-09	UBD	CpG: 80
chr6	72130209	72131020	12	6.25E-09	LINC00472; LINC01626	CpG: 138
chr8	67874858	67875646	5	6.41E-09	TCF24	CpG: 209
chr4	155662795	155664311	12	6.43E-09	LRAT	CpG: 38
chr10	125732373	125732842	5	6.74E-09	CHST15	CpG: 53
chr6	28411030	28411423	9	6.99E-09	ZSCAN23	CpG: 40
chr1	207818395	207818493	5	7.11E-09	CR1L	CpG: 53
chr13	36871646	36872346	13	7.97E-09	CCDC169; CCDC169; CCDC169-SOHLH2	CpG: 33
chr19	58570419	58570995	8	7.99E-09	ZNF135	CpG: 118
chr7	43153023	43153595	6	8.13E-09	HECW1	CpG: 124

chr4	62066211	62067393	7	8.47E-09	MIR548AG1	CpG: 212
chr12	114846849	114847641	8	8.55E-09	TBX5-AS1	CpG: 119
chr6	108485837	108487506	13	8.72E-09	NR2E1	CpG: 318
chr18	18822579	18823509	8	9.09E-09	GREB1L	CpG: 195
chr3	192232468	192233231	5	9.14E-09	FGF12	CpG: 92
chr13	37004536	37005582	27	9.24E-09	CCNA1	CpG: 103
chr16	49311483	49312560	10	9.39E-09	CBLN1	CpG: 61
chr6	84419189	84419360	9	9.39E-09	SNAP91	CpG: 185
chr1	75600124	75601276	10	9.62E-09	LHX8	CpG: 55
chr10	49731375	49732003	5	9.67E-09	ARHGAP22	CpG: 112
chr13	93878969	93879769	9	1.01E-08	GPC6	CpG: 133
chr10	22541995	22542726	5	1.05E-08	LOC100130992	CpG: 184
chr20	57089933	57090317	9	1.10E-08	APCDD1L	CpG: 71
chr8	49647579	49648297	9	1.15E-08	EFCAB1	CpG: 37
chr5	1875611	1876397	7	1.15E-08	IRX4	CpG: 306
chr20	61050885	61051915	24	1.16E-08	GATA5	CpG: 247
chr18	7117680	7118122	5	1.21E-08	LAMA1	CpG: 136
chr1	67217673	67218505	13	1.23E-08	TCTEX1D1	CpG: 23
chr1	58715499	58716033	7	1.26E-08	DAB1	CpG: 87
chr6	32115964	32116591	13	1.27E-08	PRRT1	CpG: 56
chr19	56904442	56905152	12	1.40E-08	ZNF582; ZNF582; ZNF582-AS1	CpG: 69
chr8	53477881	53478102	5	1.43E-08	FAM150A	CpG: 119
chr5	37838290	37838871	6	1.45E-08	GDNF	CpG: 308
chr20	37434158	37434552	8	1.48E-08	PPP1R16B	CpG: 126
chr20	58179847	58180616	5	1.52E-08	PHACTR3	CpG: 98

chr11	104034227	104035204	9	1.55E-08	PDGFD	CpG: 54
chr14	60975811	60976285	8	1.61E-08	SIX6	CpG: 178
chr1	147782116	147782558	5	1.63E-08	NBPF8	CpG: 32
chr4	134070039	134070819	10	1.64E-08	LOC101927359; PCDH10; PCDH10	CpG: 90
chr2	220361467	220362242	7	1.66E-08	LOC100996693	CpG: 141
chr6	33131745	33134078	32	1.78E-08	COL11A2	CpG: 47
chr1	164290179	164290833	8	1.81E-08	PBX1	CpG: 34
chr19	11784246	11785248	12	1.90E-08	ZNF833P	CpG: 20
chr14	24803679	24804342	10	1.98E-08	ADCY4	CpG: 59
chr12	106978920	106979874	9	2.00E-08	LOC100287944; RFX4	CpG: 140
chr1	119542159	119542629	6	2.02E-08	LOC105378933	CpG: 32
chr7	28997403	28998540	11	2.02E-08	TRIL	CpG: 318
chr1	217309568	217310572	7	2.07E-08	ESRRG	CpG: 46
chr1	75138964	75139958	11	2.16E-08	ERICH3	CpG: 40
chr19	15580310	15580721	6	2.28E-08	PGLYRP2	CpG: 60
chr13	43148017	43149043	9	2.28E-08	TNFSF11	CpG: 83
chr1	108023249	108023482	5	2.35E-08	NTNG1	CpG: 22
chr3	62363466	62364390	8	2.36E-08	FEZF2	CpG: 23
chr7	96625955	96627253	8	2.43E-08	DLX6-AS1	CpG: 30
chr13	37248078	37248838	6	2.53E-08	SERTM1	CpG: 60
chr10	134755862	134756707	8	2.59E-08	CFAP46	CpG: 112
chr8	109095264	109096151	12	2.67E-08	RSPO2	CpG: 118
chr18	25757202	25757710	5	2.79E-08	CDH2	CpG: 223
chr2	145274874	145275441	7	2.82E-08	ZEB2	CpG: 148
chr2	185462269	185463218	11	2.95E-08	ZNF804A	CpG: 38

chr10	100993553	100994143	11	2.96E-08	HPSE2	CpG: 38
chr6	10390039	10390961	6	2.96E-08	TFAP2A	CpG: 16; CpG: 40
chr6	72595944	72596748	10	3.04E-08	RIMS1	CpG: 33
chr16	215410	216450	10	3.07E-08	HBM	CpG: 198
chr5	169930535	169931363	9	3.12E-08	KCNIP1	CpG: 50
chr3	62358234	62358980	6	3.23E-08	FEZF2	CpG: 162
chr8	17270604	17271535	10	3.26E-08	MTMR7	CpG: 64
chr4	96468962	96469634	9	3.26E-08	UNC5C	CpG: 20
chr16	22825621	22826243	7	3.31E-08	HS3ST2	CpG: 187
chr10	110671609	110672231	5	3.33E-08	LINC01435	CpG: 43
chr19	54023732	54024834	15	3.54E-08	ZNF331	CpG: 45; CpG: 20
chr14	57274550	57275413	8	3.59E-08	OTX2	CpG: 176
chr6	127440000	127441199	10	3.61E-08	RSPO3	CpG: 152
chr16	82659960	82660873	15	3.78E-08	CDH13	CpG: 99
chr4	144621898	144621971	5	3.86E-08	FREM3	CpG: 123
chr4	111554966	111555503	5	4.00E-08	PITX2	CpG: 46
chr5	113391125	113392117	9	4.09E-08	KCNN2	CpG: 76
chr17	27044169	27044745	5	4.12E-08	RAB34; NARR; RAB34	CpG: 95
chr6	10398494	10399046	5	4.24E-08	TFAP2A	CpG: 19
chr5	1882520	1883514	10	4.35E-08	IRX4	CpG: 470
chr1	2983926	2984869	11	4.36E-08	LINC00982	CpG: 406
chr2	213401195	213402433	7	4.50E-08	ERBB4	CpG: 132
chr4	156680007	156681047	9	4.90E-08	GUCY1B3	CpG: 119
chr20	37434950	37435716	5	5.01E-08	PPP1R16B	CpG: 126

chr5	1445354	1445593	6	5.01E-08	SLC6A3	CpG: 206
chr6	108495385	108495985	5	5.02E-08	NR2E1	CpG: 27
chr7	112726089	112726869	11	5.13E-08	GPR85	CpG: 44
chr10	93392399	93393238	13	5.24E-08	PPP1R3C	CpG: 49
chr8	97157453	97158052	6	5.38E-08	GDF6	CpG: 106
chr3	185911208	185911885	6	5.44E-08	DGKG	CpG: 69
chr19	37997294	37998171	13	5.46E-08	ZNF793-AS1; ZNF793	CpG: 26
chr1	91300215	91300559	7	5.47E-08	LOC105378853	CpG: 86
chr10	102419209	102419617	6	5.48E-08	PAX2	CpG: 46
chr3	6902337	6903327	8	5.61E-08	GRM7	CpG: 77
chr5	146257347	146258785	19	5.61E-08	PPP2R2B	CpG: 97
chr8	72470352	72471153	6	5.77E-08	EYA1	CpG: 82
chr4	122686038	122686667	10	5.90E-08	TMEM155; PP12613; TMEM155	CpG: 83
chr12	128752040	128752586	5	5.97E-08	TMEM132C	CpG: 184
chr5	45695241	45696455	7	6.24E-08	HCN1	CpG: 98
chr14	95234658	95235489	9	6.35E-08	GSC	CpG: 191
chr1	114696350	114697113	11	6.54E-08	SYT6	CpG: 155; CpG: 25
chr7	30721797	30722643	11	6.63E-08	CRHR2	CpG: 107
chr6	33943520	33943929	5	6.86E-08	MIR1275	CpG: 33
chr2	87017419	87018104	6	7.25E-08	CD8A	CpG: 170
chr19	21657529	21658001	6	7.33E-08	LINC00664	CpG: 39
chr7	44143774	44144434	6	7.49E-08	AEBP1	CpG: 24
chr10	118892211	118893430	9	7.66E-08	VAX1	CpG: 36
chr7	49815117	49815751	6	7.72E-08	VWC2	CpG: 251

chr4	5712581	5713100	8	8.90E-08	EVC	CpG: 63
chr5	16179633	16180419	15	9.10E-08	Mar-11	CpG: 134
chr7	45961078	45961569	16	9.17E-08	IGFBP3	CpG: 139
chr19	15343786	15344364	5	9.39E-08	EPHX3	CpG: 33
chr16	28074384	28075289	9	9.40E-08	GSG1L	CpG: 145
chr15	83348985	83349864	7	9.50E-08	AP3B2; CPEB1-AS1	CpG: 87
chr1	46951291	46951683	6	9.52E-08	DMBX1	CpG: 48
chr18	35145983	35147532	28	9.75E-08	CELF4	CpG: 196
chr10	102483843	102484499	6	9.84E-08	PAX2	CpG: 30
chr10	7452242	7452950	5	1.02E-07	SFMBT2	CpG: 418
chr7	96649956	96650668	8	1.03E-07	DLX5	CpG: 109
chr5	54518469	54519307	7	1.05E-07	MCIDAS	CpG: 54
chr6	110678920	110679566	5	1.05E-07	METTL24	CpG: 45; CpG: 25
chr18	11148510	11149068	9	1.07E-07	PIEZO2	CpG: 164
chr14	60952097	60952945	8	1.08E-07	C14orf39	CpG: 77
chr3	69591633	69592090	6	1.08E-07	FRMD4B	CpG: 99
chr8	31496644	31497082	5	1.11E-07	NRG1	CpG: 143
chr12	72665282	72665880	5	1.15E-07	TRHDE-AS1	CpG: 175
chr10	125650957	125651621	9	1.17E-07	CPXM2	CpG: 55
chr17	42733527	42733994	9	1.18E-07	MEIOC	CpG: 101
chr1	91192162	91192803	6	1.21E-07	BARHL2	CpG: 147
chr6	110797233	110798022	6	1.29E-07	SLC22A16	CpG: 97
chr22	19511707	19512320	6	1.32E-07	CLDN5	CpG: 147
chr19	12203029	12203644	6	1.32E-07	ZNF788	CpG: 41
chr1	150254280	150255241	10	1.36E-07	CIART	CpG: 22

chr14	62583787	62584600	12	1.40E-07	LINC00643	CpG: 54
chr5	63257499	63257941	10	1.44E-07	HTR1A	CpG: 100
chr8	54164051	54164442	8	1.46E-07	OPRK1	CpG: 114
chr12	72666976	72667707	10	1.46E-07	TRHDE; TRHDE-AS1	CpG: 175
chr8	99961376	99962347	7	1.51E-07	OSR2	CpG: 73
chr12	120032374	120033155	5	1.51E-07	TMEM233	CpG: 118
chr14	57263993	57264919	7	1.56E-07	OTX2	CpG: 83
chr17	1961109	1961778	5	1.58E-07	HIC1	CpG: 764
chr10	28034669	28035208	9	1.63E-07	MKX; MKX-AS1	CpG: 398
chr19	15121204	15121596	8	1.64E-07	CCDC105; SLC1A6	CpG: 69
chr11	71954982	71955599	11	1.65E-07	PHOX2A	CpG: 89
chr19	52900882	52901307	11	1.73E-07	ZNF528; ZNF528-AS1	CpG: 24; CpG: 15
chr19	34112229	34113010	9	1.81E-07	CHST8	CpG: 232
chr10	104170217	104170718	5	1.84E-07	PSD	CpG: 55
chr11	65600739	65601427	14	1.87E-07	SNX32	CpG: 53
chr2	5835853	5836366	6	1.97E-07	SOX11	CpG: 138
chr6	30080189	30080782	7	2.02E-07	TRIM31-AS1; TRIM31; TRIM31	CpG: 15
chr11	123300839	123301171	5	2.03E-07	MIR4493	CpG: 108
chr13	26625089	26625396	7	2.04E-07	SHISA2	CpG: 192
chr3	132756721	132757741	14	2.05E-07	TMEM108	CpG: 83
chr1	14925078	14925524	6	2.08E-07	KAZN	CpG: 119
chr2	71126606	71128191	24	2.10E-07	VAX2	CpG: 185
chr2	137522211	137522960	5	2.12E-07	THSD7B	CpG: 107
chr6	168078772	168079427	6	2.17E-07	LOC401286	CpG: 19

chr13	112760898	112761184	5	2.21E-07	LINC00403	CpG: 23
chr19	45737483	45738115	9	2.31E-07	EXOC3L2	CpG: 87
chr5	170735754	170736572	11	2.34E-07	TLX3	CpG: 411
chr6	55443757	55444488	10	2.41E-07	HMGCLL1	CpG: 36
chr14	70316621	70317239	7	2.41E-07	SMOC1	CpG: 32
chr4	55991418	55992458	10	2.46E-07	KDR	CpG: 71
chr6	30227729	30228254	16	2.65E-07	HLA-L; HCG17	CpG: 83
chr11	31839314	31839947	9	2.65E-07	PAX6; DKFZp686K1684	CpG: 35
chr7	19812327	19813149	11	2.65E-07	TMEM196	CpG: 29
chr9	970297	971046	6	2.71E-07	DMRT1	CpG: 260
chr16	67977865	67978450	8	2.71E-07	SLC12A4; LCAT	CpG: 93
chr10	131758112	131758786	6	2.73E-07	EBF3	CpG: 181
chr5	63986130	63987010	6	2.79E-07	FAM159B	CpG: 78
chr4	174439478	174440103	6	2.83E-07	HAND2	CpG: 36
chr11	32454718	32455735	12	2.93E-07	WT1	CpG: 188
chr11	133938676	133939265	7	2.97E-07	JAM3	CpG: 77
chr1	77332991	77333229	8	2.97E-07	ST6GALNAC5	CpG: 112
chr20	31780882	31781697	8	3.13E-07	BPIFA4P	CpG: 30
chr1	99470550	99470801	5	3.18E-07	LOC100129620	CpG: 95
chr7	120968877	120969174	8	3.19E-07	WNT16	CpG: 81
chr15	48937856	48938775	7	3.27E-07	FBN1	CpG: 174
chr5	32712623	32713723	6	3.31E-07	NPR3	CpG: 298
chr3	172165696	172166517	7	3.44E-07	GHSR	CpG: 107
chr1	37500195	37500790	6	3.60E-07	GRIK3	CpG: 232

chr3	126113640	126114064	9	3.62E-07	CFAP100; CCDC37-AS1	CpG: 47
chr18	5543271	5544169	10	3.66E-07	EPB41L3	CpG: 82
chr19	13616871	13617584	9	3.87E-07	CACNA1A	CpG: 56
chr13	110960177	110960563	5	4.09E-07	COL4A2	CpG: 185
chr14	26674046	26674563	7	4.16E-07	NOVA1	CpG: 30
chr22	33453893	33454632	11	4.41E-07	SYN3	CpG: 68
chr22	45404843	45406130	15	4.89E-07	PHF21B	CpG: 273
chr1	228651848	228652581	7	5.01E-07	MIR4666A	CpG: 84
chr1	18958084	18958924	6	5.12E-07	PAX7	CpG: 205
chr8	57358590	57359414	9	5.18E-07	LOC101929415; PENK	CpG: 138
chr10	1779432	1780003	9	5.20E-07	ADARB2	CpG: 95
chr13	36705446	36705622	5	5.35E-07	DCLK1	CpG: 65
chr5	1886828	1887583	5	5.40E-07	CTD-2194D22.4; IRX4	CpG: 470
chr5	135265753	135266135	6	5.73E-07	FBXL21	CpG: 80
chr6	151814892	151815526	7	5.76E-07	CCDC170	CpG: 64
chr18	904851	905611	14	5.88E-07	ADCYAP1	CpG: 330
chr7	155260310	155261295	6	5.90E-07	EN2	CpG: 196
chr19	37406932	37407486	12	6.04E-07	ZNF829; ZNF568	CpG: 48
chr12	41085901	41086308	8	6.14E-07	CNTN1	CpG: 47
chr3	45077254	45078075	8	6.19E-07	CLEC3B	CpG: 54
chr6	28602543	28603230	29	6.25E-07	LINC00533	CpG: 32
chr7	116962950	116963802	7	6.33E-07	WNT2	CpG: 148
chr11	115630531	115631452	12	6.36E-07	LINC00900	CpG: 68
chr7	101005832	101006573	11	6.54E-07	COL26A1	CpG: 153

chr1	181287301	181287967	7	6.56E-07	GM140	CpG: 52
chr8	85094437	85094923	5	6.62E-07	RALYL	CpG: 49
chr4	154713235	154713789	6	6.63E-07	SFRP2	CpG: 52
chr8	65291287	65292321	12	6.71E-07	MIR124-2HG	CpG: 68
chr14	42075406	42075975	7	6.80E-07	LRFN5	CpG: 58
chr6	99272559	99273679	13	6.92E-07	POU3F2	CpG: 50
chr15	79724517	79724802	6	7.07E-07	KIAA1024	CpG: 165
chr6	32063991	32064258	12	7.08E-07	TNXB	CpG: 150
chr7	35294430	35295198	6	7.27E-07	TBX20	CpG: 124
chr1	170630397	170630734	5	7.34E-07	PRRX1	CpG: 38
chr3	184301211	184301826	5	7.40E-07	EPHB3	CpG: 92
chr10	125425583	125426048	6	7.44E-07	GPR26	CpG: 126
chr6	27647713	27648605	9	7.77E-07	LINC01012	CpG: 19
chr7	79081791	79083056	11	7.78E-07	MAGI2-AS3; MAGI2	CpG: 26; CpG: 70; CpG: 20
chr11	128693677	128694679	9	7.83E-07	FLI1	CpG: 56
chr19	57862242	57862713	10	7.83E-07	ZNF304	CpG: 58
chr5	1885174	1885950	6	7.89E-07	IRX4	CpG: 470
chr2	164592477	164592854	6	7.94E-07	FIGN	CpG: 57
chr8	79427993	79428725	8	8.08E-07	PKIA	CpG: 46
chr13	25620954	25621735	6	8.16E-07	PABPC3	CpG: 52
chr1	217310887	217311177	6	8.29E-07	ESRRG	CpG: 46
chr2	1748112	1748876	6	8.37E-07	PXDN	CpG: 198

chr15	83316223	83316911	6	8.82E-07	CPEB1-AS1; CPEB1; CPEB1	CpG: 218
chr7	601283	601828	5	8.97E-07	LOC101927000; PRKAR1B	CpG: 49
chr1	177139876	177140680	9	9.29E-07	BRINP2	CpG: 23
chr12	81471311	81472177	13	9.34E-07	ACSS3	CpG: 43
chr11	101454317	101454823	8	9.58E-07	TRPC6	CpG: 114
chr7	155174508	155175340	6	9.69E-07	BLACE	CpG: 78
chr5	115298511	115299088	5	9.74E-07	LVRN	CpG: 72
chr18	43913147	43914264	10	9.86E-07	RNF165	CpG: 103
chr2	31456741	31457287	12	1.01E-06	EHD3	CpG: 79
chr1	242688081	242688681	7	1.01E-06	PLD5	CpG: 60
chr5	127873283	127873980	5	1.03E-06	FBN2	CpG: 186
chr6	30418508	30419493	18	1.03E-06	HLA-E	CpG: 54
chr12	66122804	66123380	5	1.05E-06	RPSAP52	CpG: 56
chr6	137816475	137817512	7	1.10E-06	OLIG3	CpG: 61
chr20	54579814	54580196	5	1.13E-06	CBLN4	CpG: 173
chr4	66535145	66535732	9	1.14E-06	EPHA5-AS1; EPHA5; EPHA5	CpG: 36
chr11	8102450	8103101	6	1.21E-06	TUB	CpG: 62
chr7	49812836	49813111	7	1.22E-06	VWC2	CpG: 251
chr17	27038664	27039132	8	1.29E-06	PROCA1	CpG: 79
chr3	27770852	27772201	12	1.33E-06	EOMES	CpG: 24
chr17	50235737	50236765	7	1.35E-06	CA10	CpG: 109
chr5	92931173	92932120	7	1.35E-06	MIR548AO	CpG: 27
chr7	136553243	136553884	12	1.42E-06	CHRM2	CpG: 154
chr21	31311797	31312643	7	1.46E-06	GRIK1	CpG: 61

chr6	10422139	10422636	5	1.48E-06	TFAP2A	CpG: 26
chr7	92237896	92238364	6	1.49E-06	CDK6	CpG: 77
chr3	170302640	170303721	9	1.52E-06	SLC7A14	CpG: 21; CpG: 24
chr6	31733799	31734580	18	1.57E-06	VWA7	CpG: 68
chr6	152128024	152129036	21	1.58E-06	ESR1	CpG: 89
chr2	45170725	45171818	7	1.60E-06	SIX3	CpG: 171
chr3	148415443	148416019	6	1.61E-06	AGTR1	CpG: 56
chr17	35296481	35297309	5	1.65E-06	LHX1	CpG: 597
chr3	96533258	96534005	9	1.67E-06	EPHA6	CpG: 174
chr19	57018614	57019373	9	1.75E-06	ZNF471	CpG: 77
chr2	220299484	220300242	11	1.83E-06	SPEG	CpG: 77
chr15	89910632	89911298	6	1.85E-06	MIR9-3HG; MIR9-3	CpG: 131
chr2	229045958	229046785	7	1.85E-06	SPHKAP	CpG: 45
chr5	115297488	115298079	6	1.87E-06	LVRN	CpG: 33
chr2	119916431	119916686	7	1.90E-06	C1QL2	CpG: 229
chr15	74421523	74421935	12	1.90E-06	LOC283731; ISLR2	CpG: 216
chr7	37488162	37489005	15	1.93E-06	ELMO1	CpG: 135
chr5	33936171	33936752	13	1.95E-06	RXFP3	CpG: 166
chr12	41086680	41087252	6	1.98E-06	CNTN1	CpG: 47
chr11	69631981	69633313	9	1.99E-06	FGF3	CpG: 263
chr8	15397637	15398333	8	2.01E-06	TUSC3	CpG: 67
chr5	174158984	174159705	5	2.09E-06	MSX2	CpG: 88
chr4	156128904	156130016	11	2.10E-06	NPY2R	CpG: 67
chr10	134599783	134600600	20	2.12E-06	NKX6-2	CpG: 569
chr21	28339262	28339907	7	2.13E-06	ADAMTS5	CpG: 193

chr1	63795279	63795934	6	2.20E-06	MIR6068	CpG: 66
chr8	42623686	42623946	6	2.21E-06	CHRNA6	CpG: 55
chr17	73583839	73584617	9	2.29E-06	MYO15B	CpG: 221
chr4	111549880	111550666	6	2.38E-06	PITX2	CpG: 22
chr8	97505391	97505868	8	2.43E-06	SDC2	CpG: 165
chr1	91190005	91190891	8	2.44E-06	BARHL2	CpG: 147
chr10	102893925	102894639	6	2.46E-06	TLX1NB; TLX1	CpG: 110
chr2	30144152	30144838	8	2.57E-06	ALK	CpG: 147
chr2	119600348	119600627	5	2.58E-06	EN1	CpG: 103
chr3	63263752	63264335	8	2.59E-06	SYNPR	CpG: 20
chr20	39995539	39996039	8	2.71E-06	EMILIN3	CpG: 131
chr11	30038286	30039113	14	2.71E-06	KCNA4	CpG: 32; CpG: 16
chrX	130191788	130192734	12	2.71E-06	LINC01201; ARHGAP36	CpG: 55
chr5	79330669	79331135	7	2.72E-06	THBS4	CpG: 57
chr10	122216848	122217376	5	2.73E-06	PLPP4	CpG: 118
chr4	41258560	41259283	7	2.85E-06	UCHL1; UCHL1-AS1	CpG: 123
chr8	97169764	97170666	6	2.86E-06	GDF6	CpG: 56
chr5	131592959	131593287	9	2.86E-06	PDLIM4	CpG: 64
chr10	23983496	23983990	7	2.87E-06	KIAA1217	CpG: 139
chr7	21582444	21582834	6	3.11E-06	DNAH11	CpG: 84
chr18	31020418	31021065	8	3.22E-06	CCDC178	CpG: 41
chr20	21378045	21378560	10	3.33E-06	NKX2-4	CpG: 213
chr8	109094032	109094880	7	3.70E-06	RSPO2	CpG: 118
chr5	177412586	177413021	5	3.82E-06	PROP1	CpG: 17

chr3	121902442	121902622	7	3.83E-06	CASR	CpG: 75
chr3	13323567	13324132	6	3.84E-06	NUP210	CpG: 136
chr5	76249502	76250351	5	3.87E-06	CRHBP	CpG: 104
chr1	221054201	221055097	7	3.96E-06	HLX	CpG: 24
chr17	35300448	35300874	5	3.99E-06	LHX1	CpG: 597
chr8	16884197	16884576	6	4.06E-06	MICU3	CpG: 103
chr7	155164995	155165891	6	4.11E-06	BLACE	CpG: 277
chr3	13114570	13115406	8	4.30E-06	IQSEC1	CpG: 70
chr11	69633659	69634334	6	4.34E-06	FGF3	CpG: 263
chr3	26663805	26664741	9	4.43E-06	LRRC3B	CpG: 71
chr5	137474700	137475288	7	4.47E-06	NME5	CpG: 23
chr8	97507561	97507958	5	4.56E-06	SDC2	CpG: 165
chr2	240168836	240169280	6	4.57E-06	HDAC4	CpG: 22
chr13	38443634	38444227	9	4.65E-06	TRPC4	CpG: 45
chr5	146889027	146889701	6	4.69E-06	DPYSL3	CpG: 102
chr12	54333179	54333823	6	4.70E-06	HOXC13; HOXC13-AS	CpG: 88
chr7	96647021	96647483	6	4.75E-06	DLX5	CpG: 60
chr11	17740887	17741243	5	4.79E-06	MYOD1	CpG: 233
chr1	49242359	49242934	8	4.82E-06	BEND5; AGBL4;BEND5	CpG: 38
chr2	127413363	127414108	8	4.89E-06	GYPC	CpG: 54
chr14	42076492	42077327	6	5.11E-06	LRFN5	CpG: 58
chr10	134598908	134599372	5	5.33E-06	NKX6-2	CpG: 569
chr8	132052702	132052942	6	5.42E-06	ADCY8	CpG: 208
chr2	115419537	115420260	6	5.42E-06	DPP10	CpG: 51
chrX	153637261	153637827	10	5.43E-06	DNASE1L1	CpG: 93

chr7	30951272	30951801	11	5.47E-06	AQP1	CpG: 64
chr5	155108085	155109126	6	5.58E-06	SGCD	CpG: 138
chr5	38258007	38258884	11	5.73E-06	EGFLAM	CpG: 108
chr7	84815088	84815543	5	5.75E-06	SEMA3D	CpG: 94
chr12	103889166	103890274	13	5.97E-06	C12orf42; C12orf42; LOC105369945	CpG: 79
chr4	5891833	5892628	7	6.19E-06	CRMP1	CpG: 35
chr9	124461008	124461640	5	6.22E-06	DAB2IP	CpG: 38
chr20	42544648	42545099	10	6.46E-06	TOX2	CpG: 172
chr8	26722496	26722965	5	6.55E-06	ADRA1A	CpG: 202
chr11	32452038	32452839	8	6.60E-06	WT1	CpG: 53
chr7	100463206	100463759	5	6.61E-06	SLC12A9	CpG: 32
chr4	7194841	7195466	5	6.62E-06	SORCS2	CpG: 99
chr5	59189375	59189934	10	6.74E-06	PDE4D	CpG: 68
chr6	36808207	36808894	5	6.87E-06	CPNE5	CpG: 80
chr6	85484287	85484863	5	6.91E-06	TBX18	CpG: 188
chr8	23567310	23567941	7	6.99E-06	NKX2-6	CpG: 43
chr5	128300696	128301488	12	7.13E-06	SLC27A6	CpG: 44
chr13	100623942	100624293	8	7.24E-06	ZIC5	CpG: 391
chr5	140857474	140858107	8	7.30E-06	PCDHGA8; PCDHGA1; PCDHGB7; PCDHGC3; PCDHGA11; PCDHGB5; PCDHGA9; PCDHGA7; PCDHGA10; PCDHGA5; PCDHGA2; PCDHGA3; PCDHGA12; PCDHGB6; PCDHGB4; PCDHGB3; PCDHGB2; PCDHGB1; PCDHGA6; PCDHGA4	CpG: 19
chr11	43602845	43603214	8	7.42E-06	MIR129-2	CpG: 61
chr1	203044677	203045229	5	7.42E-06	PPFIA4	CpG: 67
chr3	6903921	6904640	5	7.61E-06	GRM7	CpG: 40
chr8	11561283	11561724	10	7.69E-06	GATA4	CpG: 221

chr3	159756648	159757071	6	7.73E-06	IL12A-AS1	CpG: 35
chr5	127874463	127874825	5	7.84E-06	FBN2	CpG: 186
chr16	56228114	56228901	12	8.17E-06	GNAO1; DKFZP434H168	CpG: 91
chr1	53308654	53309261	5	8.35E-06	ZYG11A	CpG: 112
chr1	13910555	13910796	6	8.49E-06	PDPN	CpG: 59
chr5	145713607	145714097	7	8.54E-06	POU4F3	CpG: 22
chr15	84115433	84116151	12	8.70E-06	SH3GL3	CpG: 95
chr5	1931251	1931781	5	8.76E-06	CTD-2194D22.4	CpG: 110
chr16	875410	876216	7	8.92E-06	PRR25	CpG: 37
chr6	28641704	28642394	11	9.25E-06	LINC00533	CpG: 32
chr1	53098630	53099352	7	9.29E-06	FAM159A	CpG: 70
chr6	55039232	55039622	5	9.30E-06	HCRTR2	CpG: 15
chr18	44777736	44778236	5	9.36E-06	SKOR2	CpG: 45
chr11	12030088	12030289	6	9.40E-06	DKK3	CpG: 103
chr2	54086854	54087552	13	9.59E-06	GPR75;GPR75-ASB3	CpG: 60
chr10	102821427	102821848	6	9.68E-06	KAZALD1	CpG: 173
chr9	122131895	122132261	6	1.03E-05	BRINP1	CpG: 104
chr14	95239381	95239751	5	1.03E-05	GSC	CpG: 35
chr2	45028604	45030072	11	1.06E-05	CAMKMT	CpG: 32; CpG: 38
chr20	61885850	61886313	7	1.06E-05	NKAIN4; FLJ16779	CpG: 156
chr7	3341524	3342048	6	1.09E-05	SDK1	CpG: 24
chr4	81106663	81107185	5	1.11E-05	PRDM8	CpG: 65
chr8	70946891	70947440	5	1.20E-05	PRDM14	CpG: 39
chr5	82767297	82767968	7	1.23E-05	VCAN	CpG: 63
chr15	47476133	47476482	6	1.27E-05	SEMA6D	CpG: 95

chr16	54968727	54969146	6	1.28E-05	IRX5	CpG: 453
chr6	31894990	31895598	7	1.29E-05	C2	CpG: 124
chr6	123317124	123317714	9	1.29E-05	CLVS2	CpG: 59
chr8	91803357	91804110	8	1.32E-05	NECAB1	CpG: 48
chr10	102896102	102896869	6	1.32E-05	TLX1; TLX1NB	CpG: 30
chr7	79083408	79083509	5	1.33E-05	MAGI2-AS3	CpG: 70
chr8	72756656	72757004	8	1.34E-05	MSC; MSC-AS1	CpG: 88
chr3	147136904	147137503	7	1.36E-05	LOC440982	CpG: 37
chr2	182548750	182549408	7	1.37E-05	NEUROD1	CpG: 94
chr13	103046499	103047287	7	1.39E-05	FGF14;FGF14-AS2; FGF14-IT1	CpG: 60
chr6	100050506	100051339	8	1.39E-05	PRDM13	CpG: 83
chr11	31825605	31825969	5	1.40E-05	PAX6	CpG: 81
chr18	76740088	76740284	7	1.42E-05	SALL3	CpG: 338
chr11	65360123	65360509	5	1.45E-05	KCNK7	CpG: 114
chr17	27369780	27370117	5	1.51E-05	PIPOX	CpG: 34
chr1	12123262	12123936	9	1.57E-05	TNFRSF8	CpG: 63
chr8	9760869	9761454	7	1.61E-05	MIR124-1	CpG: 70
chr8	85096835	85097429	9	1.66E-05	RALYL	CpG: 49
chr13	107028903	107029443	5	1.75E-05	LINC00460	CpG: 30
chr3	141516232	141516705	6	1.76E-05	GRK7	CpG: 43
chr2	233924713	233925275	9	1.88E-05	INPP5D	CpG: 18
chr13	95655083	95655579	5	1.90E-05	ABCC4	CpG: 46
chr14	77737495	77738080	7	1.91E-05	NGB	CpG: 116
chr6	85472949	85473344	5	1.95E-05	TBX18	CpG: 129
chr17	72352968	72353536	5	1.96E-05	BTBD17	CpG: 129

chr17	37365885	37366501	5	2.07E-05	STAC2	CpG: 54
chr2	27530670	27531360	11	2.09E-05	UCN	CpG: 180
chr4	107956955	107957430	5	2.11E-05	DKK2	CpG: 62
chr11	111385338	111385778	6	2.20E-05	C11orf88	CpG: 31
chr17	75368750	75369657	10	2.21E-05	Septin9	CpG: 138
chr16	66612955	66613582	12	2.21E-05	CKLF-CMTM1; CMTM2; CMTM1	CpG: 70
chr15	35046610	35047203	7	2.23E-05	GJD2	CpG: 106
chr19	36909326	36909913	6	2.28E-05	ZFP82	CpG: 66
chr7	54609587	54610320	11	2.34E-05	VSTM2A	CpG: 39
chr17	8770908	8771331	7	2.36E-05	PIK3R6	CpG: 35
chr7	94284380	94285004	24	2.40E-05	SGCE	CpG: 112
chr8	58906736	58907387	7	2.49E-05	FAM110B	CpG: 95
chr10	82116203	82116571	9	2.60E-05	DYDC2; DYDC2; DYDC2; DYDC1	CpG: 93
chr1	3567109	3567550	7	2.73E-05	WRAP73	CpG: 284
chr16	3238937	3239638	5	2.75E-05	OR1F1	CpG: 51
chr1	119530600	119531122	6	2.82E-05	TBX15	CpG: 84
chr3	238048	238618	8	2.92E-05	CHL1	CpG: 171
chr3	62354878	62355525	7	2.94E-05	FEZF2	CpG: 16; CpG: 59
chr1	229543009	229543603	5	2.99E-05	ACTA1	CpG: 69
chr5	170743459	170744027	5	3.01E-05	TLX3	CpG: 72
chr10	22634038	22634602	11	3.29E-05	SPAG6	CpG: 102
chr3	32859099	32859587	8	3.35E-05	TRIM71	CpG: 251
chr10	8077830	8078357	7	3.36E-05	GATA3-AS1	CpG: 41
chr5	40841298	40841590	5	3.41E-05	CARD6	CpG: 53

chr2	70995349	70995459	5	3.46E-05	ADD2	CpG: 88
chr12	33591899	33592642	5	3.58E-05	SYT10	CpG: 129
chr2	394115	394572	5	3.61E-05	FAM150B	CpG: 116
chr8	9763033	9763739	7	3.62E-05	MIR124-1	CpG: 170
chr6	166422182	166422741	5	3.70E-05	LINC00602	CpG: 37
chr10	102997782	102998117	6	3.73E-05	LBX1-AS1	CpG: 40
chr1	25257505	25258082	13	3.86E-05	RUNX3	CpG: 311
chr12	104850321	104850767	9	3.91E-05	CHST11	CpG: 184
chr6	10385320	10386160	7	3.91E-05	TFAP2A	CpG: 49
chr2	242794854	242795282	6	4.05E-05	PDCD1	CpG: 19
chr6	10882927	10883376	5	4.14E-05	GCM2	CpG: 21
chr1	221057558	221057808	5	4.25E-05	HLX	CpG: 24
chr12	132663428	132663883	8	4.34E-05	GALNT9	CpG: 32
chr3	238931	239733	7	4.40E-05	CHL1	CpG: 171
chr11	32456759	32457386	10	4.44E-05	WT1-AS; WT1	CpG: 188
chr2	119615496	119616242	5	4.47E-05	EN1	CpG: 173; CpG: 48
chr19	53635967	53636398	7	4.53E-05	ZNF415	CpG: 45
chr3	157155088	157155476	5	4.62E-05	VEPH1; PTX3	CpG: 67
chrX	133305793	133306382	5	4.73E-05	MIR106A	CpG: 111
chr14	60977856	60978275	5	4.73E-05	SIX6	CpG: 178
chr12	95942761	95943114	5	4.83E-05	USP44	CpG: 134
chr15	83952345	83952875	7	4.84E-05	BNC1	CpG: 188
chr6	101846779	101847541	13	4.90E-05	GRIK2	CpG: 27
chr13	102068234	102068650	5	4.96E-05	NALCN	CpG: 99
chr19	9473565	9474128	11	5.02E-05	ZNF177; ZNF559-ZNF177	CpG: 31

chr3	44626029	44626860	11	5.07E-05	ZNF660	CpG: 35
chr19	49249932	49250561	7	5.11E-05	IZUMO1	CpG: 44
chr11	119293385	119293869	5	5.34E-05	THY1 USP2-AS1;THY1	CpG: 48
chr11	79149704	79150593	6	5.45E-05	TENM4	CpG: 300
chr5	50685201	50685716	5	5.46E-05	ISL1	CpG: 53
chr3	170137240	170137871	5	5.55E-05	CLDN11	CpG: 132
chr4	13536409	13537318	8	5.66E-05	LINC01097	CpG: 27
chr1	208083913	208084415	6	5.68E-05	CD34	CpG: 36
chr4	20255061	20255343	5	5.68E-05	SLIT2	CpG: 291
chr8	101117949	101118548	9	5.70E-05	RGS22	CpG: 88
chr6	30720080	30720491	8	5.74E-05	IER3	CpG: 161
chr14	60973462	60974039	6	5.88E-05	SIX6	CpG: 32
chr1	78956845	78957253	5	6.09E-05	PTGFR	CpG: 70
chr7	96651915	96652481	10	6.21E-05	DLX5	CpG: 20
chr6	124124620	124125371	7	6.22E-05	NKAIN2	CpG: 117
chr2	207506975	207507528	7	6.27E-05	LOC200726	CpG: 53
chr6	50817157	50817674	6	6.30E-05	TFAP2B	CpG: 39
chr6	50803820	50804174	6	6.50E-05	TFAP2B	CpG: 29
chr1	10732063	10733030	11	6.52E-05	CASZ1	CpG: 36
chr6	6004207	6004897	8	6.52E-05	NRN1	CpG: 191
chr10	7708670	7709073	7	6.71E-05	ITIH5	CpG: 21
chr3	44039204	44040206	8	6.80E-05	MIR138-1	CpG: 344
chr2	234847554	234848156	5	7.20E-05	TRPM8	CpG: 19
chr2	154729622	154730157	5	7.34E-05	GALNT13	CpG: 23
chr3	147142102	147142415	5	7.97E-05	LOC440982	CpG: 18

chr20	21494547	21494831	8	8.00E-05	NKX2-2	CpG: 671
chrX	101905837	101906288	5	8.21E-05	ARMCX5-GPRASP2	CpG: 89
chr1	101004472	101005121	5	8.40E-05	GPR88	CpG: 134
chr19	2251764	2252432	6	8.47E-05	JSRP1; AMH	CpG: 330
chr15	27111435	27112430	21	8.48E-05	GABRA5	CpG: 159
chrX	107979401	107980144	7	8.56E-05	IRS4; LOC101928358	CpG: 160
chr11	109292789	109293216	7	9.03E-05	C11orf87	CpG: 102
chr3	172167445	172167810	5	9.10E-05	GHSR	CpG: 26
chr12	129338484	129338939	5	9.77E-05	GLT1D1	CpG: 93
chr10	102995987	102996798	8	0.0001009	LBX1-AS1	CpG: 40
chr8	97171668	97172219	5	0.0001012	GDF6	CpG: 24
chr16	1030166	1030619	9	0.0001013	LMF1	CpG: 437
chr20	30060822	30061044	5	0.0001092	DEFB124	CpG: 27
chr19	21646329	21646781	5	0.0001138	LINC00664	CpG: 18
chr8	141248582	141249094	6	0.0001142	TRAPPC9	CpG: 32
chr3	62359390	62360240	9	0.0001158	FEZF2	CpG: 162
chr3	128719999	128720433	6	0.0001163	EFCC1	CpG: 139
chr11	122850140	122850741	6	0.0001171	BSX	CpG: 86
chr8	23563572	23564294	10	0.0001181	NKX2-6	CpG: 187
chr6	118228382	118228871	8	0.0001236	SLC35F1	CpG: 166
chr4	93225933	93226540	6	0.0001274	GRID2	CpG: 63
chr4	41882564	41883172	9	0.0001286	LINC00682	CpG: 34
chr2	223166909	223167463	6	0.0001295	CCDC140	CpG: 25
chr11	31819148	31819678	7	0.0001449	PAX6	CpG: 98
chr19	54926277	54926805	7	0.000146	TTYH1	CpG: 35

chr5	26669320	26669882	5	0.0001464	CDH9	CpG: 28
chr6	33135753	33136511	9	0.0001517	COL11A2	CpG: 47
chr12	99139301	99139867	6	0.000156	ANKS1B	CpG: 42
chr12	104851909	104852446	6	0.0001586	CHST11	CpG: 184
chr10	94825257	94825848	5	0.0001712	CYP26C1	CpG: 64
chr19	51071294	51071328	5	0.0001716	LRRC4B	CpG: 20
chr13	103052362	103052943	5	0.0001722	FGF14	CpG: 51
chr18	11149435	11150043	5	0.0001736	PIEZO2	CpG: 164
chr20	45279830	45280414	7	0.0001767	SLC13A3	CpG: 40
chr12	114846162	114846503	8	0.0001773	TBX5-AS1; TBX5	CpG: 119
chr11	66188019	66188745	6	0.0001852	NPAS4	CpG: 102
chr19	52956622	52957180	10	0.0001927	ZNF578	CpG: 47
chr17	78152051	78152466	5	0.0001931	CARD14	CpG: 15
chr11	45376874	45377324	5	0.0002003	LOC101928812	CpG: 25
chr16	54972573	54973385	6	0.0002086	IRX5	CpG: 179
chr10	125852693	125853232	6	0.0002089	CHST15	CpG: 209
chr1	3331987	3332264	6	0.0002108	PRDM16	CpG: 19
chr5	77268331	77268888	5	0.0002117	LOC101929154	CpG: 44
chr18	59001158	59001726	5	0.0002172	CDH20	CpG: 101
chr16	66637919	66638702	12	0.0002207	CMTM3	CpG: 126
chr12	3600019	3600317	6	0.0002248	THCAT155; PRMT8	CpG: 212
chr12	4918848	4919230	5	0.0002308	KCNA6	CpG: 138
chr7	27142200	27143046	15	0.0002344	HOXA2	CpG: 24
chr17	21280740	21281318	5	0.0002443	KCNJ12	CpG: 187
chr1	1566687	1567257	5	0.0002511	MMP23B	CpG: 470
chr8	144240945	144241434	5	0.0002568	LY6H	CpG: 132

chr4	185942086	185942746	7	0.0002626	HELT	CpG: 256
chr20	61885249	61885291	5	0.0002648	NKAIN4	CpG: 156
chr15	88799300	88799999	7	0.0002651	NTRK3-AS1; NTRK3; NTRK3	CpG: 119
chr13	36919344	36919960	7	0.000273	SPG20	CpG: 121
chr17	4689640	4689893	6	0.0002878	VMO1	CpG: 75
chr12	128850196	128850696	6	0.0003024	TMEM132C	CpG: 17
chr11	123301490	123302099	6	0.0003047	MIR4493	CpG: 108
chr4	185940923	185941625	6	0.0003057	HELT	CpG: 256
chr6	10883834	10884314	6	0.000311	GCM2	CpG: 21
chr5	87968118	87968658	5	0.0003147	LINC00461	CpG: 24
chr7	19146032	19146555	5	0.0003232	TWIST1	CpG: 34
chr11	3239953	3240399	6	0.0003282	MRGPRG; MRGPRG-AS1	CpG: 66
chr12	15475116	15475767	8	0.0003309	PTPRO	CpG: 61
chr17	74944775	74945219	5	0.0003334	MGAT5B	CpG: 17
chr7	142966	143220	5	0.0003355	LOC102723672	CpG: 49
chr12	82153008	82153464	9	0.0003366	PPFIA2	CpG: 34
chr10	1778529	1779033	5	0.0003429	ADARB2	CpG: 95
chr1	2980038	2980598	6	0.0003437	LINC00982	CpG: 126
chr7	88388555	88388646	5	0.0003533	ZNF804B	CpG: 71
chr7	45613410	45613825	5	0.0003668	ADCY1	CpG: 151
chr4	172733760	172734347	5	0.0003677	LOC441052	CpG: 117
chr15	89951787	89952499	5	0.0003712	MIR9-3HG	CpG: 58
chr10	119000638	119001529	8	0.000375	SLC18A2	CpG: 123
chr5	161274540	161275138	6	0.0003787	GABRA1	CpG: 68
chr17	74868561	74869001	5	0.0003822	MGAT5B	CpG: 218

chr3	147125712	147126244	13	0.0004167	ZIC1	CpG: 180
chr13	112547459	112548065	5	0.0004431	LINC00354	CpG: 53
chr21	32930938	32931935	6	0.0004478	TIAM1	CpG: 249
chr6	85482946	85483800	5	0.0004663	TBX18	CpG: 188
chr11	19367501	19368112	6	0.0005057	NAV2	CpG: 116
chr14	23821149	23822017	7	0.0005071	SLC22A17	CpG: 92
chr4	134069236	134069670	6	0.0005101	LOC101927359	CpG: 90
chr3	85008587	85008991	5	0.0005216	CADM2	CpG: 48
chr17	43974804	43975063	5	0.0005542	MAPT-IT1; MAPT	CpG: 302
chrX	30327423	30327819	12	0.0005624	NR0B1	CpG: 116
chr3	27765232	27765546	5	0.0005948	EOMES	CpG: 40
chr14	104690040	104690244	5	0.0006521	KIF26A	CpG: 18
chr10	28035631	28035822	5	0.0006524	MKX-AS1	CpG: 398
chr1	34630518	34631450	8	0.0007154	CSMD2	CpG: 169
chr10	25465171	25465352	6	0.0007992	GPR158; GPR158-AS1	CpG: 166
chr1	236849653	236850232	5	0.0008241	ACTN2	CpG: 84
chr20	24451428	24452036	6	0.0008448	SYNDIG1	CpG: 173
chr7	70597458	70597921	7	0.0008857	WBSCR17	CpG: 232
chr3	42306385	42306974	5	0.0009454	CCK	CpG: 94
chr4	81187601	81187906	5	0.001249	FGF5	CpG: 43
chr12	45444203	45445023	7	0.001305	DBX2	CpG: 115
chr10	130338781	130339186	5	0.001393	LINC01163	CpG: 28
chr20	30062696	30063249	7	0.001405	REM1	CpG: 27
chr11	134145889	134146324	9	0.001533	GLB1L3	CpG: 152
chr10	94820892	94821085	7	0.001618	CYP26C1	CpG: 255

chr15	74422275	74422761	11	0.001873	ISLR2	CpG: 216
chr4	175750280	175750666	9	0.002116	GLRA3	CpG: 26
chr11	20181725	20182324	9	0.002204	DBX1	CpG: 87
chr10	99790170	99790596	5	0.00222	CRTAC1	CpG: 160
chr11	109293945	109294233	6	0.002269	C11orf87	CpG: 102
chr6	152957910	152958173	7	0.00227	SYNE1	CpG: 60
chr5	180017623	180018201	5	0.002534	SCGB3A1	CpG: 176
chr17	8868669	8869213	8	0.002554	PIK3R5	CpG: 87
chr1	116380946	116381475	6	0.002677	NHLH2	CpG: 209
chrX	66765795	66766334	6	0.0029	AR	CpG: 27
chr11	94501461	94502008	5	0.003172	AMOTL1	CpG: 114
chr2	124782586	124782885	5	0.003638	CNTNAP5	CpG: 71
chr8	72917147	72917695	5	0.003748	MSC-AS1	CpG: 29
chr2	200334558	200335148	7	0.004256	SATB2; SATB2-AS1	CpG: 74
chr5	119800143	119800972	5	0.004356	PRR16	CpG: 52; CpG: 15
chr2	63283806	63284132	5	0.004399	OTX1	CpG: 26
chr20	24449734	24449916	5	0.004702	SYNDIG1	CpG: 173
chr3	181441571	181442137	5	0.004754	SOX2-OT	CpG: 47
chr8	37823979	37824529	7	0.005035	ADRB3	CpG: 137
chr2	161127261	161127827	5	0.005904	LOC100505984	CpG: 39
chrX	142723112	142723713	5	0.00617	SLITRK4	CpG: 135
chr2	163695776	163696000	5	0.006657	KCNH7	CpG: 46
chr6	71666682	71666858	5	0.007235	B3GAT2	CpG: 181
chr2	63274615	63275178	5	0.008145	LOC100132215	CpG: 354
chr1	107683532	107684059	6	0.008488	NTNG1	CpG: 134

chr14	92790097	92790711	5	0.01163	SLC24A4	CpG: 131
chr16	86546785	86547386	6	0.01236	FOXF1	CpG: 21
chr11	637032	637175	5	0.01455	DRD4	CpG: 278
chr7	32110145	32110650	5	0.01587	PDE1C	CpG: 84
chr1	13910138	13910224	6	0.01677	PDPN	CpG: 59
chr15	92936818	92937360	5	0.01742	ST8SIA2	CpG: 177
chr2	88751621	88752322	7	0.01812	FOXI3	CpG: 120
chr2	121223740	121224009	6	0.01835	LINC01101	CpG: 34
chr4	126237371	126237597	6	0.02693	FAT4	CpG: 230
chr12	96883202	96883596	5	0.036	CFAP54	CpG: 45
chr10	128994030	128994644	6	0.04364	FAM196A; DOCK1	CpG: 172
chr15	92937735	92938295	5	0.04917	ST8SIA2	CpG: 177

Supplementary Table 5. GO enrichment analysis – Biological category

Term	P value	Adjusted P value
embryonic morphogenesis (GO:0048598)	1.7132E-17	5.72552E-14
behaviour (GO:0007610)	4.80146E-15	8.02324E-12
pattern specification process (GO:0007389)	1.02545E-14	1.14235E-11
synaptic transmission (GO:0007268)	2.04466E-13	1.13888E-10
cell-cell adhesion (GO:0098609)	5.09521E-14	3.40564E-11
cell-cell adhesion via plasma-membrane adhesion molecules (GO:0098742)	4.44464E-14	3.40564E-11
regulation of neuron differentiation (GO:0045664)	2.62233E-12	1.09548E-09
neuron differentiation (GO:0030182)	2.43946E-12	1.09548E-09
homophilic cell adhesion via plasma membrane adhesion molecules (GO:0007156)	3.94219E-12	1.46387E-09
single-organism behaviour (GO:0044708)	3.60916E-11	1.20618E-08
regionalization (GO:0003002)	1.37505E-10	4.17766E-08
organ morphogenesis (GO:0009887)	8.93567E-10	1.99087E-07
central nervous system neuron differentiation (GO:0021953)	4.8044E-10	1.33803E-07
embryonic organ morphogenesis (GO:0048562)	6.33846E-10	1.62947E-07
locomotory behaviour (GO:0007626)	3.00094E-09	6.26821E-07
cell fate specification (GO:0001708)	8.08533E-10	1.93008E-07
tissue morphogenesis (GO:0048729)	8.93821E-09	1.65953E-06
forelimb morphogenesis (GO:0035136)	1.76139E-08	2.67572E-06
negative regulation of cell development (GO:0010721)	1.17763E-08	2.07138E-06
neuron migration (GO:0001764)	7.58469E-09	1.49106E-06
cell fate commitment (GO:0045165)	1.6157E-08	2.57127E-06

appendage morphogenesis (GO:0035107)	2.29374E-08	3.19403E-06
limb morphogenesis (GO:0035108)	2.29374E-08	3.19403E-06
behavioural fear response (GO:0001662)	1.22961E-07	1.3256E-05
behavioural defense response (GO:0002209)	1.22961E-07	1.3256E-05
negative regulation of nervous system development (GO:0051961)	4.47994E-08	5.98878E-06
multicellular organismal response to stress (GO:0033555)	1.39448E-08	2.33017E-06
hindlimb morphogenesis (GO:0035137)	8.15769E-08	1.04858E-05
morphogenesis of an epithelium (GO:0002009)	9.38606E-08	1.157E-05
anterior/posterior pattern specification (GO:0009952)	9.69364E-08	1.157E-05
negative regulation of neurogenesis (GO:0050768)	1.05065E-07	1.21079E-05
fear response (GO:0042596)	3.92016E-07	3.44768E-05
positive regulation of nervous system development (GO:0051962)	2.98233E-07	3.05004E-05
cell fate determination (GO:0001709)	3.11388E-07	3.05004E-05
signal transduction involved in regulation of gene expression (GO:0023019)	6.02994E-07	4.79811E-05
embryonic appendage morphogenesis (GO:0035113)	3.37677E-07	3.05004E-05
embryonic limb morphogenesis (GO:0030326)	3.37677E-07	3.05004E-05
morphogenesis of a branching epithelium (GO:0061138)	3.21392E-07	3.05004E-05
adenylate cyclase-modulating G-protein coupled receptor signaling pathway (GO:0007188)	3.19789E-07	3.05004E-05
adult behaviour (GO:0030534)	5.97081E-07	4.79811E-05
embryonic forelimb morphogenesis (GO:0035115)	1.7115E-06	0.000111787
morphogenesis of a branching structure (GO:0001763)	6.73236E-07	5.23245E-05
palate development (GO:0060021)	5.43109E-07	4.65403E-05
branching morphogenesis of an epithelial tube (GO:0048754)	5.97081E-07	4.79811E-05
extracellular matrix organization (GO:0030198)	1.73936E-06	0.000111787
negative regulation of neuron differentiation (GO:0045665)	1.0751E-06	7.81085E-05

tube development (GO:0035295)	1.40714E-06	9.40533E-05
extracellular structure organization (GO:0043062)	1.83937E-06	0.000115984
neuron development (GO:0048666)	1.1828E-06	8.41048E-05
embryonic skeletal system morphogenesis (GO:0048704)	8.27848E-07	6.28788E-05
nervous system development (GO:0007399)	2.03006E-06	0.000123354
muscle organ development (GO:0007517)	9.28862E-07	6.89835E-05
sensory perception (GO:0007600)	2.70182E-06	0.000161241
G-protein coupled receptor signaling pathway, coupled to cyclic nucleotide second messenger (GO:0007187)	1.8803E-06	0.000116369
muscle structure development (GO:0061061)	1.31039E-06	9.09979E-05
cell differentiation in spinal cord (GO:0021515)	3.99096E-06	0.000229962
neurogenesis (GO:0022008)	4.86299E-06	0.00027546
adenylate cyclase-inhibiting G-protein coupled receptor signaling pathway (GO:0007193)	1.3342E-06	9.09979E-05
embryonic hindlimb morphogenesis (GO:0035116)	4.96874E-06	0.000276759
axon guidance (GO:0007411)	6.90918E-06	0.000360789
neuron projection guidance (GO:0097485)	6.90918E-06	0.000360789
regulation of organ morphogenesis (GO:2000027)	5.13471E-06	0.000281315
sympathetic nervous system development (GO:0048485)	7.88033E-06	0.00040517
positive regulation of neurogenesis (GO:0050769)	8.2482E-06	0.000417659
positive regulation of cell development (GO:0010720)	1.09239E-05	0.000536875
neuron fate specification (GO:0048665)	6.86565E-06	0.000360789
skeletal system morphogenesis (GO:0048705)	3.91051E-06	0.000229279
adult locomotory behaviour (GO:0008344)	9.18401E-06	0.000458104
neural crest cell migration (GO:0001755)	1.42802E-05	0.000681778
regulation of cell morphogenesis involved in differentiation (GO:0010769)	2.11006E-05	0.000979418
regulation of membrane potential (GO:0042391)	2.24333E-05	0.001027017

skeletal system development (GO:0001501)	1.3918E-05	0.000674116
epithelial cell differentiation (GO:0030855)	2.54267E-05	0.001133014
regulation of organ formation (GO:0003156)	2.80573E-05	0.001186931
pituitary gland development (GO:0021983)	2.80573E-05	0.001186931
cellular component morphogenesis (GO:0032989)	3.28287E-05	0.001371418
regulation of transmembrane receptor protein serine/threonine kinase signaling pathway (GO:0090092)	2.63132E-05	0.001142405
positive regulation of neuron differentiation (GO:0045666)	2.42282E-05	0.001094197
cranial nerve development (GO:0021545)	4.19925E-05	0.001690832
regulation of neuron projection development (GO:0010975)	3.73187E-05	0.001520964
feeding behaviour (GO:0007631)	3.44487E-05	0.001421328
neuron-neuron synaptic transmission (GO:0007270)	1.99371E-05	0.000938449
inner ear morphogenesis (GO:0042472)	2.63211E-05	0.001142405
regulation of developmental growth (GO:0048638)	4.98756E-05	0.001984335
middle ear morphogenesis (GO:0042474)	5.31752E-05	0.002090725
response to steroid hormone (GO:0048545)	0.000105577	0.003675392
regulation of transmembrane transport (GO:0034762)	8.52796E-05	0.003131916
gland development (GO:0048732)	8.20531E-05	0.00307709
epithelium development (GO:0060429)	8.9379E-05	0.003211878
ameboidal-type cell migration (GO:0001667)	7.41025E-05	0.00281421
cell morphogenesis (GO:0000902)	8.62213E-05	0.003132081
positive regulation of heart growth (GO:0060421)	8.2866E-05	0.00307709
regulation of ion transmembrane transport (GO:0034765)	0.000110179	0.003796065
positive regulation of organ growth (GO:0046622)	0.000104421	0.003673436
central nervous system development (GO:0007417)	7.08495E-05	0.002721596
dorsal/ventral pattern formation (GO:0009953)	5.71215E-05	0.002219767

regulation of stem cell proliferation (GO:0072091)	0.000123508	0.004211872
growth (GO:0040007)	0.000219614	0.006924069
regulation of BMP signaling pathway (GO:0030510)	0.000128105	0.004324526
positive regulation of developmental growth (GO:0048639)	0.000102682	0.003650664
positive regulation of growth (GO:0045927)	0.000266953	0.008078519
kidney epithelium development (GO:0072073)	0.000157199	0.005253601
potassium ion transport (GO:0006813)	0.00019337	0.006397681
regulation of DNA binding (GO:0051101)	0.000195287	0.006397681
kidney development (GO:0001822)	0.000227967	0.007054305
regulation of peptide secretion (GO:0002791)	0.000318379	0.009252382
positive regulation of cardiac muscle cell proliferation (GO:0060045)	0.000197176	0.006397681
neuromuscular process (GO:0050905)	0.000238779	0.007321108
axis specification (GO:0009798)	0.000216109	0.00687845
negative regulation of Wnt signaling pathway (GO:0030178)	0.000308412	0.00904134
regulation of Wnt signaling pathway (GO:0030111)	0.00041175	0.010835193
digestive tract development (GO:0048565)	0.000268317	0.008078519
cellular response to lipid (GO:0071396)	0.000617915	0.014048116
nerve development (GO:0021675)	0.000494563	0.011890863
regulation of neural precursor cell proliferation (GO:2000177)	0.000221816	0.006928117
regulation of cell morphogenesis (GO:0022604)	0.000506993	0.01210265
midbrain development (GO:0030901)	0.000307993	0.00904134
positive regulation of DNA binding (GO:0043388)	0.000362491	0.010266472
positive regulation of amine transport (GO:0051954)	0.000372959	0.010386915
cardiac ventricle morphogenesis (GO:0003208)	0.000452905	0.011174413
branching involved in ureteric bud morphogenesis (GO:0001658)	0.000349106	0.010057851
tube formation (GO:0035148)	0.000399616	0.010599347

cell morphogenesis involved in differentiation (GO:0000904)	0.000438756	0.011174413
visual perception (GO:0007601)	0.000378209	0.010425362
positive regulation of cardiac muscle tissue growth (GO:0055023)	0.000304483	0.00904134
regulation of axonogenesis (GO:0050770)	0.000399616	0.010599347
potassium ion transmembrane transport (GO:0071805)	0.000383698	0.010425362
metanephros development (GO:0001656)	0.000494563	0.011890863
spinal cord motor neuron differentiation (GO:0021522)	0.000372959	0.010386915
cellular potassium ion transport (GO:0071804)	0.000383698	0.010425362
neuron fate commitment (GO:0048663)	0.000424441	0.011046489
sensory perception of light stimulus (GO:0050953)	0.000447556	0.011174413
positive regulation of cell adhesion (GO:0045785)	0.000687152	0.015009551
neuropeptide signaling pathway (GO:0007218)	0.000399616	0.010599347
blood vessel morphogenesis (GO:0048514)	0.000562137	0.013046261
regulation of embryonic development (GO:0045995)	0.000473637	0.011553966
regulation of peptide hormone secretion (GO:0090276)	0.000713541	0.015384866
fluid transport (GO:0042044)	0.000454734	0.011174413
regulation of transcription regulatory region DNA binding (GO:2000677)	0.000452905	0.011174413
cellular response to steroid hormone stimulus (GO:0071383)	0.000770084	0.015502591
regulation of system process (GO:0044057)	0.001178409	0.02030022
regulation of hormone secretion (GO:0046883)	0.000925567	0.017761122
negative regulation of transmembrane receptor protein serine/threonine kinase signaling pathway (GO:0090101)	0.0007565	0.015502591
cyclic nucleotide metabolic process (GO:0009187)	0.000429696	0.011046489
regulation of osteoblast differentiation (GO:0045667)	0.000701398	0.015221254
response to corticosteroid (GO:0031960)	0.000808587	0.016085111
neuron projection morphogenesis (GO:0048812)	0.000863438	0.016874914

positive regulation of epithelial cell proliferation (GO:0050679)	0.001039532	0.019300641
regulation of postsynaptic membrane potential (GO:0060078)	0.000429696	0.011046489
regulation of cell projection organization (GO:0031344)	0.001170177	0.020262856
regulation of amine transport (GO:0051952)	0.000536256	0.012532632
cardiac chamber morphogenesis (GO:0003206)	0.000774666	0.015502591
ionotropic glutamate receptor signaling pathway (GO:0035235)	0.000452905	0.011174413
stem cell differentiation (GO:0048863)	0.000663563	0.014686276
cell differentiation in hindbrain (GO:0021533)	0.000652383	0.014632651
negative regulation of cAMP metabolic process (GO:0030815)	0.000573611	0.013130187
regulation of epithelial cell proliferation (GO:0050678)	0.001412474	0.023139651
synaptic transmission, glutamatergic (GO:0035249)	0.000573611	0.013130187
neural tube closure (GO:0001843)	0.000678503	0.01491814
regulation of striated muscle tissue development (GO:0016202)	0.001089263	0.019627288
positive regulation of embryonic development (GO:0040019)	0.000652383	0.014632651
developmental growth (GO:0048589)	0.001166756	0.020262856
negative regulation of cyclic nucleotide metabolic process (GO:0030800)	0.000662376	0.014686276
regulation of muscle tissue development (GO:1901861)	0.001168599	0.020262856
tube closure (GO:0060606)	0.000743739	0.015465269
regulation of binding (GO:0051098)	0.001255902	0.02119811
circulatory system process (GO:0003013)	0.001228283	0.020943484
developmental growth involved in morphogenesis (GO:0060560)	0.000889742	0.017287893
positive regulation of stem cell proliferation (GO:2000648)	0.00081454	0.01610764
muscle tissue development (GO:0060537)	0.001142905	0.020103095
regulation of muscle organ development (GO:0048634)	0.001252667	0.02119811
adult walking behaviour (GO:0007628)	0.000761688	0.015502591
regulation of smoothened signalling pathway (GO:0008589)	0.000735901	0.015465269

gonad development (GO:0008406)	0.001089263	0.019627288
outflow tract morphogenesis (GO:0003151)	0.000995443	0.018689718
aging (GO:0007568)	0.001548017	0.02487246
response to glucocorticoid (GO:0051384)	0.001228283	0.020943484
phospholipase C-activating G-protein coupled receptor signaling pathway (GO:0007200)	0.000935355	0.017761122
columnar/cuboidal epithelial cell differentiation (GO:0002065)	0.000835104	0.016417164
developmental induction (GO:0031128)	0.000774666	0.015502591
brain development (GO:0007420)	0.00163934	0.025842799
endoderm formation (GO:0001706)	0.000354722	0.010132304
glutamate receptor signaling pathway (GO:0007215)	0.000995443	0.018689718
cognition (GO:0050890)	0.001907182	0.028771434
cell-cell signaling involved in cell fate commitment (GO:0045168)	0.000913913	0.017654905
striated muscle tissue development (GO:0014706)	0.001474646	0.023808056
embryonic cranial skeleton morphogenesis (GO:0048701)	0.001071651	0.019570814
regulation of excitatory postsynaptic membrane potential (GO:0060079)	0.000935355	0.017761122
regulation of peptide transport (GO:0090087)	0.002277122	0.032203671
second-messenger-mediated signaling (GO:0019932)	0.001579397	0.025255237
regulation of ossification (GO:0030278)	0.002008912	0.029972243
regulation of gliogenesis (GO:0014013)	0.001092363	0.019627288
neurotransmitter transport (GO:0006836)	0.001329004	0.022319255
positive regulation of cardiac muscle tissue development (GO:0055025)	0.001071651	0.019570814
regulation of heart growth (GO:0060420)	0.001131718	0.020011651
response to alcohol (GO:0097305)	0.002724808	0.0359769
cardiac septum morphogenesis (GO:0060411)	0.001447895	0.023506985
regulation of cardiac muscle cell proliferation (GO:0060043)	0.001071651	0.019570814
regulation of cell adhesion (GO:0030155)	0.002925203	0.037600107

calcium ion homeostasis (GO:0055074)	0.002724808	0.0359769
enteric nervous system development (GO:0048484)	0.000199755	0.006419059
negative regulation of cell morphogenesis involved in differentiation (GO:0010771)	0.001866844	0.028359051
negative regulation of transport (GO:0051051)	0.003401837	0.042901664
regulation of gastrulation (GO:0010470)	0.001448964	0.023506985
regulation of vasculature development (GO:1901342)	0.002671431	0.0359769
membrane depolarization (GO:0051899)	0.002186163	0.03188907
regulation of angiogenesis (GO:0045765)	0.00272243	0.0359769
learning or memory (GO:0007611)	0.002799177	0.036149633
mesonephric tubule development (GO:0072164)	0.001629824	0.025814558
signal release (GO:0023061)	0.002403624	0.033610503
negative regulation of canonical Wnt signaling pathway (GO:0090090)	0.002339284	0.032848266
negative regulation of purine nucleotide metabolic process (GO:1900543)	0.001721906	0.026518939
positive regulation of cell-substrate adhesion (GO:0010811)	0.002704379	0.0359769
mesonephric epithelium development (GO:0072163)	0.001829047	0.027911756
axonogenesis (GO:0007409)	0.002669987	0.0359769
negative regulation of gliogenesis (GO:0014014)	0.001671845	0.02586715
regulation of cell growth (GO:0001558)	0.003647827	0.045151991
neurotransmitter secretion (GO:0007269)	0.00159062	0.025313588
positive regulation of ion transport (GO:0043270)	0.00380721	0.046606949
response to pain (GO:0048265)	0.001671845	0.02586715
regulation of synaptic transmission (GO:0050804)	0.004078595	0.049386465
response to fibroblast growth factor (GO:0071774)	0.003987647	0.048460788
positive regulation of muscle organ development (GO:0048636)	0.002046649	0.030131728
negative regulation of nucleotide metabolic process (GO:0045980)	0.002221102	0.031973599
positive regulation of striated muscle tissue development (GO:0045844)	0.002046649	0.030131728

cellular response to organic cyclic compound (GO:0071407)	0.005115282	0.057754304
cyclic purine nucleotide metabolic process (GO:0052652)	0.001919817	0.028771434
negative regulation of cAMP biosynthetic process (GO:0030818)	0.001671845	0.02586715
nitric oxide mediated signal transduction (GO:0007263)	0.000738126	0.015465269
regulation of synaptic transmission, glutamatergic (GO:0051966)	0.002283743	0.032203671
response to alkaloid (GO:0043279)	0.003038036	0.038900826
cellular response to hormone stimulus (GO:0032870)	0.005493761	0.060000484
cyclic nucleotide biosynthetic process (GO:0009190)	0.00219464	0.03188907
positive regulation of muscle tissue development (GO:1901863)	0.002283743	0.032203671
negative regulation of cyclic nucleotide biosynthetic process (GO:0030803)	0.001919817	0.028771434
activation of protein kinase A activity (GO:0034199)	0.000738126	0.015465269
rhythmic process (GO:0048511)	0.005347711	0.058789639
formation of primary germ layer (GO:0001704)	0.002046649	0.030131728
negative regulation of secretion (GO:0051048)	0.004340393	0.051427935
divalent inorganic cation homeostasis (GO:0072507)	0.005115282	0.057754304
regulation of canonical Wnt signaling pathway (GO:0060828)	0.004340393	0.051427935
response to cocaine (GO:0042220)	0.00219464	0.03188907
regulation of cardiac muscle tissue development (GO:0055024)	0.002541464	0.034667648
stem cell proliferation (GO:0072089)	0.002541464	0.034667648
digestive tract morphogenesis (GO:0048546)	0.00112612	0.020011651
cell-matrix adhesion (GO:0007160)	0.004354909	0.051427935
negative regulation of nucleotide biosynthetic process (GO:0030809)	0.002498102	0.034356618
negative regulation of purine nucleotide biosynthetic process (GO:1900372)	0.002498102	0.034356618
Wnt signaling pathway (GO:0016055)	0.005731808	0.062193834
neural crest cell development (GO:0014032)	0.00112612	0.020011651
regulation of cardiac muscle tissue growth (GO:0055021)	0.002498102	0.034356618

anion transmembrane transport (GO:0098656)	0.00472011	0.054963786
muscle cell differentiation (GO:0042692)	0.004938096	0.057104216
single organismal cell-cell adhesion (GO:0016337)	0.005958903	0.063991171
sensory perception of pain (GO:0019233)	0.003123444	0.039690302
positive regulation of angiogenesis (GO:0045766)	0.004644407	0.054271353
sensory organ development (GO:0007423)	0.005268696	0.058593999
regulation of lipase activity (GO:0060191)	0.005855247	0.063327621
regulation of glial cell differentiation (GO:0045685)	0.003123444	0.039690302
single organism cell adhesion (GO:0098602)	0.006923424	0.069068901
positive regulation of peptide secretion (GO:0002793)	0.003955147	0.048241249
regulation of heart morphogenesis (GO:2000826)	0.001368616	0.022531595
cellular response to estradiol stimulus (GO:0071392)	0.001368616	0.022531595
cell-substrate adhesion (GO:0031589)	0.00591971	0.063818288
embryonic digestive tract morphogenesis (GO:0048557)	0.00164753	0.025849978
cAMP metabolic process (GO:0046058)	0.002498102	0.034356618
cellular response to fibroblast growth factor stimulus (GO:0044344)	0.007748532	0.075718114
sensory perception of sound (GO:0007605)	0.006537663	0.068065011
reproductive structure development (GO:0048608)	0.008556455	0.081936029
cell growth (GO:0016049)	0.006326206	0.066276425
response to cAMP (GO:0051591)	0.006677545	0.068665713
regulation of synapse assembly (GO:0051963)	0.005023774	0.057301888
response to drug (GO:0042493)	0.010060453	0.090870364
positive regulation of lipase activity (GO:0060193)	0.006677545	0.068665713
sensory perception of mechanical stimulus (GO:0050954)	0.007676419	0.075233405
renal system process (GO:0003014)	0.005435766	0.059561743
cell projection morphogenesis (GO:0048858)	0.009712379	0.088443512

memory (GO:0007613)	0.006677545	0.068665713
embryonic digit morphogenesis (GO:0042733)	0.005022791	0.057301888
regulation of cell-substrate adhesion (GO:0010810)	0.009216165	0.084849655
heart development (GO:0007507)	0.008868558	0.083254838
regulation of neurotransmitter levels (GO:0001505)	0.006677545	0.068665713
negative regulation of BMP signaling pathway (GO:0030514)	0.004509519	0.052880045
mesenchyme development (GO:0060485)	0.005022791	0.057301888
ureteric bud development (GO:0001657)	0.004509519	0.052880045
response to purine-containing compound (GO:0014074)	0.009911005	0.090007005
regulation of insulin secretion (GO:0050796)	0.009681595	0.088404072
water transport (GO:0006833)	0.005023774	0.057301888
mesenchymal cell development (GO:0014031)	0.00273433	0.0359769
hormone-mediated signaling pathway (GO:0009755)	0.008082202	0.078519532
adrenal gland development (GO:0030325)	0.00273433	0.0359769
regulation of sequence-specific DNA binding transcription factor activity (GO:0051090)	0.011779369	0.098911187
regulation of cAMP metabolic process (GO:0030814)	0.009434234	0.086618713
regulation of organ growth (GO:0046620)	0.006798634	0.069068901

Supplementary Table 6 GO enrichment analysis – Cellular category

Term	P value	Adjusted P value
synapse part (GO:0044456)	1.9776E-09	6.24921E-07
integral component of plasma membrane (GO:0005887)	3.36109E-08	5.31052E-06
extracellular matrix (GO:0031012)	7.55151E-08	7.95426E-06
dendrite (GO:0030425)	1.09876E-06	8.57239E-05
proteinaceous extracellular matrix (GO:0005578)	1.35639E-06	8.57239E-05
ion channel complex (GO:0034702)	1.38185E-05	0.000727772
postsynaptic membrane (GO:0045211)	2.01469E-05	0.000909487
transmembrane transporter complex (GO:1902495)	2.42697E-05	0.000958654
transporter complex (GO:1990351)	3.16854E-05	0.00111251
synaptic membrane (GO:0097060)	5.41966E-05	0.001712611
axon part (GO:0033267)	8.26454E-05	0.002374176
synapse (GO:0045202)	0.000125062	0.00329329
receptor complex (GO:0043235)	0.000209504	0.00509256
transcription factor complex (GO:0005667)	0.000430107	0.009263376
extracellular region (GO:0005576)	0.001272455	0.022338662
axon (GO:0030424)	0.000554144	0.01094434
ionotropic glutamate receptor complex (GO:0008328)	0.000439717	0.009263376
presynaptic membrane (GO:0042734)	0.001143319	0.021252288
extracellular matrix part (GO:0044420)	0.001732199	0.028809205
neuronal cell body (GO:0043025)	0.002542347	0.040169085
cell body (GO:0044297)	0.002842581	0.040829795

basement membrane (GO:0005604)	0.005881631	0.071484435
postsynaptic density (GO:0014069)	0.007459839	0.084189613
perikaryon (GO:0043204)	0.007427013	0.084189613
synaptic vesicle (GO:0008021)	0.010616733	0.104840238
chloride channel complex (GO:0034707)	0.008851019	0.091158777
axon terminus (GO:0043679)	0.007806262	0.085061341
extracellular space (GO:0005615)	0.015357726	0.138658327
interstitial matrix (GO:0005614)	0.002767292	0.040829795

Bibliography

Aaltonen, L. A., Peltomaki, P., Leach, F. S., Sistonen, P., Pylkkanen, L., Mecklin, J. P., … et, al. (1993). Clues to the pathogenesis of familial colorectal cancer. *Science*, *260*(5109), 812 LP-816. Retrieved from http://science.sciencemag.org/content/260/5109/812.abstract

Abe, M., Watanabe, N., McDonell, N., Takato, T., Ohira, M., Nakagawara, A., & Ushijima, T. (2008). Identification of genes targeted by CpG island methylator phenotype in neuroblastomas, and their possible integrative involvement in poor prognosis. *Oncology*, *74*. https://doi.org/10.1159/000139124

Adler, D. G., Gostout, C. J., Sorbi, D., Burgart, L. J., Wang, L., & Harmsen, W. S. (2002). Endoscopic identification and quantification of aberrant crypt foci in the human colon. *Gastrointest Endosc*, *56*. https://doi.org/10.1016/S0016-5107(02)70113-2

Antelo, M., Balaguer, F., Shia, J., Shen, Y., Hur, K., Moreira, L., … Goel, A. (2012). A high degree of LINE-1 hypomethylation is a unique feature of early-onset colorectal cancer. *PLoS One*, *7*. https://doi.org/10.1371/journal.pone.0045357

Apgar, B., Brotzman, G., & Spitzer, M. (2008). Colposcopy: Principles and practice. Retrieved from http://books.google.com/books?hl=en&lr=&id=XvxpzpjpGIAC&oi=fnd&pg=PT3&dq=Colposcopy+Principles+and+Practice&ots=OCQC66e3su&sig=xe0rS1RF3KepUjFhI8XAk2hMSlg%5Cnhttp://books.google.com/books?hl=en&lr=&id=XvxpzpjpGIAC&oi=fnd&pg=PT3&dq=Colposcopy:+Principles+a

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., … Sherlock, G. (2000). Gene Ontology: tool for the unification of biology. *Nature Genetics*, *25*(1), 25–29. https://doi.org/10.1038/75556

Ashktorab, H., Rahi, H., Wansley, D., Varma, S., Shokrani, B., Lee, E., … Brim, H. (2013). Toward a comprehensive and systematic methylome signature in colorectal cancers. *Epigenetics*, *8*(8), 807–815. https://doi.org/10.4161/epi.25497

Ashour, N., Angulo, J. C., Andrés, G., Alelú, R., González-Corpas, A., Toledo, M. V., … Ropero, S. (2014). A DNA hypermethylation profile reveals new potential biomarkers for prostate cancer diagnosis and prognosis. *The Prostate*, *74*(12), 1171–1182. https://doi.org/10.1002/pros.22833

Astler, V. B., & Coller, F. A. (1954). The Prognostic Significance of Direct Extension of Carcinoma of the Colon and Rectum . *Annals of Surgery*, *139*(6), 846–851. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1609522/

Atkin, W. S., Edwards, R., Kralj-Hans, I., Wooldrage, K., Hart, A. R., Northover, J. M., … Cuzick, J. (2010). Once-only flexible sigmoidoscopy screening in prevention of colorectal cancer: a multicentre randomised controlled trial. *Lancet*, *375*. https://doi.org/10.1016/S0140-6736(10)60551-X

Ausch, C., Kim, Y. H., Tsuchiya, K. D., Dzieciatkowski, S., Washington, M. K., Paraskeva, C., … Grady, W. M. (2009). Comparative analysis of PCR-based biomarker assay methods for colorectal polyp detection from fecal DNA. *Clin Chem*, *55*. https://doi.org/10.1373/clinchem.2008.122937

Avner, P., & Heard, E. (2001). X-chromosome inactivation: counting, choice and initiation. *Nat Rev Genet*, *2*(1), 59–67. https://doi.org/10.1038/35047580

Bae, J. M., Kim, J. H., Cho, N.-Y., Kim, T.-Y., & Kang, G. H. (2013). Prognostic implication of the CpG island methylator phenotype in colorectal cancers

depends on tumour location. *British Journal of Cancer*, *109*(4), 1004–12. https://doi.org/10.1038/bjc.2013.430

Bahrami, A., Amerizadeh, F., Hassanian, S. M., ShahidSales, S., Khazaei, M., Maftouh, M., … Avan, A. (2017). Genetic variants as potential predictive biomarkers in advanced colorectal cancer patients treated with oxaliplatin-based chemotherapy. *Journal of Cellular Physiology*. https://doi.org/10.1002/jcp.25966

Bartel, D. P. (2009). MicroRNAs: Target Recognition and Regulatory Functions. *Cell*, *136*(2), 215–233. https://doi.org/10.1016/j.cell.2009.01.002

Battersby, N. J., How, P., Moran, B., Stelzner, S., West, N. P., Branagan, G., … Brown, G. (2015). Prospective Validation of a Low Rectal Cancer Magnetic Resonance Imaging Staging System and Development of a Local Recurrence Risk Stratification Model: The MERCURY II Study. *Annals of Surgery*, 4–5. https://doi.org/10.1097/SLA.0000000000001193

Baylin, S. B., Esteller, M., Rountree, M. R., Bachman, K. E., Schuebel, K., & Herman, J. G. (2001). Aberrant patterns of DNA methylation, chromatin formation and gene expression in cancer. *Hum Mol Genet*, *10*. https://doi.org/10.1093/hmg/10.7.687

Beets-Tan, R. G. H., Beets, G. L., Vliegen, R. F. A., Kessels, A. G. H., Van Boven, H., De Bruine, A., … Van Engelshoven, J. M. A. (2001). Accuracy of magnetic resonance imaging in prediction of tumour-free resection margin in rectal cancer surgery. *Lancet*, *357*(9255), 497–504. https://doi.org/10.1016/S0140-6736(00)04040-X

Beggs, A. D., Dilworth, M. P., Domingo, E., Midgley, R., Kerr, D., Tomlinson, I. P. M., & Middleton, G. W. (2015). Methylation changes in the TFAP2E promoter region are associated with BRAF mutation and poorer overall & disease free survival in colorectal cancer. *Oncoscience*, *2*(5), 508–516. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4468337/

Beggs, A. D., Jones, A., El-Bahwary, M., Abulafi, M., Hodgson, S. V, & Tomlinson, I. P. M. (2013). Whole-genome methylation analysis of benign and malignant colorectal tumours. *The Journal of Pathology*, *229*(5), 697–704. https://doi.org/10.1002/path.4132

Bell, A., Bell, D., Weber, R. S., & El-Naggar, A. K. (2011). CpG Island Methylation Profiling in Human Salivary Gland Adenoid Cystic Carcinoma: DNA Methylation in Salivary Gland ACC. *Cancer*, *117*(13), 2898–2909. https://doi.org/10.1002/cncr.25818

Belshaw, N. J., Elliott, G. O., Foxall, R. J., Dainty, J. R., Pal, N., Coupe, A., … Johnson, I. T. (2008). Profiling CpG island field methylation in both morphologically normal and neoplastic human colonic mucosa. *British Journal of Cancer*, *99*(1), 136–142. https://doi.org/10.1038/sj.bjc.6604432

Belshaw, N. J., Pal, N., Tapp, H. S., Dainty, J. R., Lewis, M. P. N., Williams, M. R., … Johnson, I. T. (2010). Patterns of DNA methylation in individual colonic crypts reveal aging and cancer-related field defects in the morphologically normal mucosa. *Carcinogenesis*, *31*(6), 1158–1163. https://doi.org/10.1093/carcin/bgq077

Benard, a, van de Velde, C. J. H., Lessard, L., Putter, H., Takeshima, L., Kuppen, P. J. K., & Hoon, D. S. B. (2013). Epigenetic status of LINE-1 predicts clinical outcome in early-stage rectal cancer. *British Journal of Cancer*, *109*(12), 3073–83. https://doi.org/10.1038/bjc.2013.654

Bendas, G., & Borsig, L. (2012). Cancer cell adhesion and metastasis: Selectins, integrins, and the inhibitory potential of heparins. *International Journal of Cell Biology*. https://doi.org/10.1155/2012/676731

Berg, M., Nordgaard, O., Kørner, H., Oltedal, S., Smaaland, R., Søreide, J. A., & Søreide, K. (2015). Molecular Subtypes in Stage II-III Colon Cancer Defined by Genomic Instability: Early Recurrence-Risk Associated with a High Copy-Number Variation and Loss of RUNX3 and CDKN2A. *PLoS ONE*, *10*(4), e0122391. https://doi.org/10.1371/journal.pone.0122391

Berger, S. L. (2007). The complex language of chromatin regulation during transcription. *Nature*, *447*(7143), 407–412. https://doi.org/10.1038/nature05915

Bernet, A., Mazelin, L., Coissieux, M. M., Gadot, N., Ackerman, S. L., Scoazec, J. Y., & Mehlen, P. (2007). Inactivation of the UNC5C Netrin-1 Receptor Is Associated With Tumor Progression in Colorectal Malignancies. *Gastroenterology*, *133*(6), 1840–1848. https://doi.org/10.1053/j.gastro.2007.08.009

Bertrand, F. E., Angus, C. W., Partis, W. J., & Sigounas, G. (2012). Developmental pathways in colon cancer: Crosstalk between WNT, BMP, Hedgehog and Notch. *Cell Cycle*, *11*(23), 4344–4351. https://doi.org/10.4161/cc.22134

Bettoni, F., Masotti, C., Habr-Gama, A., Correa, B. R., Gama-Rodrigues, J., Vianna, M. R., … Perez, R. O. (2016). Intratumoral Genetic Heterogeneity in Rectal Cancer: Are Single Biopsies representative of the entirety of the tumor? *Annals of Surgery*. https://doi.org/10.1097/SLA.0000000000001937

Bibikova, M., Le, J., Barnes, B., Saedinia-melnyk, S., Zhou, L., Shen, R., & Gunderson, K. L. (2009). Genome-wide DNA methylation profiling using infinium(R) assay. *Epigenomics*, *1*(1), 177–200.

Biondo, S., Fraccalvieri, D., Golda, T., Frago, R., Trenti, L., & Kreisler, E. (2016). Update on advances and controversy in rectal cancer treatment. *Techniques in Coloproctology*, *20*(3), 145–152. https://doi.org/10.1007/s10151-015-1418-y

Birbeck, K. F., Macklin, C. P., Tiffin, N. J., Parsons, W., Dixon, M. F., Mapstone, N. P., … Quirke, P. (2002). Rates of Circumferential Resection Margin Involvement Vary Between Surgeons and Predict Outcomes in Rectal Cancer Surgery. *Annals of Surgery*, *235*(4), 449–457. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1422458/

Bird, a P. (1986). CpG-rich islands and the function of DNA methylation. *Nature*, *321*(6067), 209–213. https://doi.org/10.1038/321209a0

Bird, A. P., & Wolffe, A. P. (2017). Methylation-Induced Repression— Belts, Braces, and Chromatin. *Cell*, *99*(5), 451–454. https://doi.org/10.1016/S0092-8674(00)81532-9

Bird, R. P., & Good, C. K. (2000). The significance of aberrant crypt foci in understanding the pathogenesis of colon cancer. *Toxicol Lett*, *112–113*. https://doi.org/10.1016/S0378-4274(99)00261-1

Blausen.Com. (2014). Medical Gallery of Blausen Medical 2014.

Blomqvist, L., & Glimelius, B. (2008). The “good”, the “bad”, and the “ugly” rectal cancers. *Acta Oncologica*, *47*(1), 5–8. https://doi.org/10.1080/02841860701802585

Boland CR, G. A. (2010). Microsatellite instability in colorectal cancer. *Gastroenterology*, *138*, 2073–2087. https://doi.org/10.1038/nrclinonc.2009.237

Borinstein, S. C., Conerly, M., Dzieciatkowski, S., Biswas, S., Washington, M. K., Trobridge, P., … Grady, W. M. (2010). Aberrant DNA methylation occurs in colon neoplasms arising in the azoxymethane colon cancer

model. *Molecular Carcinogenesis*, *49*(1), 94–103. https://doi.org/10.1002/mc.20581
Bosch, L. J. W., Oort, F. A., Neerincx, M., Khalid-de Bakker, C. A. J., Terhaar Sive Droste, J. S., Melotte, V., … Meijer, G. A. (2012). DNA methylation of phosphatase and actin regulator 3 detects colorectal cancer in stool and complements FIT. *Cancer Prevention Research*, *5*(3), 464–472. https://doi.org/10.1158/1940-6207.CAPR-11-0315
Bottger, A., & Spruce, B. A. (1995). Proenkephalin is a nuclear protein responsive to growth arrest and differentiation signals. *J Cell Biol*, *130*. https://doi.org/10.1083/jcb.130.6.1251
Boyle, P., & Leon, M. E. (2002). Epidemiology of colorectal cancer. *Br Med Bull*, *64*. https://doi.org/10.1093/bmb/64.1.1
Brabletz, T. (2012). EMT and MET in Metastasis: Where Are the Cancer Stem Cells? *Cancer Cell*, *22*(6), 699–701. https://doi.org/10.1016/j.ccr.2012.11.009
Brenner, H., Kloor, M., & Pox, C. P. (2014). Colorectal cancer. *The Lancet*, *383*(9927), 1490–1502. https://doi.org/https://doi.org/10.1016/S0140-6736(13)61649-9
Bretthauer, M. (2010). Evidence for colorectal cancer screening. *Best Pract Res Clin Gastroenterol*, *24*. https://doi.org/10.1016/j.bpg.2010.06.005
Bretthauer, M., Gondal, G., Larsen, K., Carlsen, E., Eide, T. J., Grotmol, T., … Hoff, G. (2002). Design, organization and management of a controlled population screening study for detection of colorectal neoplasia: attendance rates in the NORCCAP study (Norwegian Colorectal Cancer Prevention). *Scand J Gastroenterol*, *37*. https://doi.org/10.1080/00365520252903125
Brown, G., & Daniels, I. R. (2005). Preoperative Staging of Rectal Cancer: The MERCURY Research Project BT - Rectal Cancer Treatment. In M. W. Büchler, J. Weitz, B. Ulrich, & R. J. Heald (Eds.) (pp. 58–74). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/3-540-27449-9_8
Buczacki, S. J. A., & Davies, R. J. (2014). The confounding effects of tumour heterogeneity and cellular plasticity on personalized surgical management of colorectal cancer. *Colorectal Disease*, *16*(5), 329–331. https://doi.org/10.1111/codi.12625
Byun, T., Karimi, M., Marsh, J. L., Milovanovic, T., Lin, F., & Holcombe, R. F. (2005). Expression of secreted Wnt antagonists in gastrointestinal tissues: potential role in stem cell homeostasis. *J Clin Pathol*, *58*. https://doi.org/10.1136/jcp.2004.018598
Cancer Genome Atlas Network, {fname}, Cancer, T., & Atlas, G. (2012). Comprehensive molecular characterization of human colon and rectal cancer. *Nature*, *487*(7407), 330–7. https://doi.org/10.1038/nature11252
Carr NJ, S. L., & Neiderau, C. (2000). Aberrrant crypt foci. Pathology and genetics: Tumors of the digestive system BT - WHO Classification of Tumours Series. Lyon: IARC Press.
Cebrat, M., Strzadala, L., & Kisielow, P. (2004). Wnt inhibitory factor-1: a candidate for a new player in tumorigenesis of intestinal epithelial cells. *Cancer Lett*, *206*. https://doi.org/10.1016/j.canlet.2003.10.024
Chan, A. O., & Rashid, A. (2006). CpG island methylation in precursors of gastrointestinal malignancies. *Curr Mol Med*, *6*. https://doi.org/10.2174/156652406777435417

Chan, A. O., Broaddus, R. R., Houlihan, P. S., Issa, J. J., Hamilton, S. R., & Rashid, A. (2002). CpG Island Methylation in Aberrant Crypt Foci of the Colorectum. *Public Health*, *160*(5), 1823–1830.

Chang, E., Park, D. I., Kim, Y. J., Kim, B. K., Park, J. H., Kim, H. J., … Kim, Y. H. (2010). Detection of colorectal neoplasm using promoter methylation of ITGA4, SFRP2, and p16 in stool samples: a preliminary report in Korean patients. *Hepatogastroenterology*, *57*.

Chang, W.-L., Lai, W.-W., Kuo, I.-Y., Lin, C.-Y., Lu, P.-J., Sheu, B.-S., & Wang, Y.-C. (2016). A six-CpG panel with DNA methylation biomarkers predicting treatment response of chemoradiation in esophageal squamous cell carcinoma. *Journal of Gastroenterology*. https://doi.org/10.1007/s00535-016-1265-2

Chen, C., Sun, P., Rong, J., Weng, H.-W., Dai, Q., & Ye, S. (2015). Short Course Radiation in the Treatment of Localized Rectal cancer: A Systematic Review and Meta-Analysis. *Scientific Reports*, *5*, 10953. https://doi.org/10.1038/srep10953

Chen, L.-F. (2012). Tumor suppressor function of RUNX3 in breast cancer. *Journal of Cellular Biochemistry*, *113*(5), 1470–1477. https://doi.org/10.1002/jcb.24074

Chen, Q., Yao, Y., Xu, H., Chen, Y., Gu, M., Cai, Z., & Wang, Z. (2016). SPOCK1 promotes tumor growth and metastasis in human prostate cancer. *Drug Design, Development and Therapy*, *10*, 2311–2321. https://doi.org/10.2147/DDDT.S91321

Chen, W. D., Han, Z. J., Skoletsky, J., Olson, J., Sah, J., Myeroff, L., … Markowitz, S. D. (2005). Detection in fecal DNA of colon cancer-specific methylation of the nonexpressed vimentin gene. *J Natl Cancer Inst*, *97*. https://doi.org/10.1093/jnci/dji204

Chung, W., Bondaruk, J., Jelinek, J., Lotan, Y., Liang, S., Czerniak, B., & Issa, J. P. (2011). Detection of bladder cancer using novel DNA methylation biomarkers in urine sediments. *Cancer Epidemiol Biomarkers*, *20*. https://doi.org/10.1158/1055-9965.EPI-11-0067

Chung, W., Kwabi-Addo, B., Ittmann, M., Jelinek, J., Shen, L., Yu, Y., & Issa, J. P. (2008). Identification of novel tumor markers in prostate, colon and breast cancer by unbiased methylation profiling. *PLoS ONE*, *3*. https://doi.org/10.1371/journal.pone.0002079

Clerc, L., Jooste, V., Lejeune, C., Schmitt, B., Arveux, P., & Quantin, C. (2008). Cost of care of colorectal cancers according to health care patterns and stage at diagnosis in France. *Eur J Health Econ*, *9*. https://doi.org/10.1007/s10198-007-0083-0

Compton, C. C. (2003). Colorectal Carcinoma: Diagnostic, Prognostic, and Molecular Features. *Mod Pathol*, *16*(4), 376–388. Retrieved from http://dx.doi.org/10.1097/01.MP.0000062859.46942.93

CONSORTIUM, G. O. (2015). GENE ONTOLOGY Website. Retrieved May 10, 2017, from http://geneontology.org

Coppedè, F. (2014). Epigenetic biomarkers of colorectal cancer: Focus on DNA methylation. *Cancer Letters*, *342*(2), 238–247. https://doi.org/https://doi.org/10.1016/j.canlet.2011.12.030

Costello, J. F., Fruhwald, M. C., Smiraglia, D. J., Rush, L. J., Robertson, G. P., Gao, X., … Plass, C. (2000). Aberrant CpG-island methylation has non-random and tumour-type-specific patterns. *Nat Genet*, *24*(2), 132–138. https://doi.org/10.1038/72785

Costello, J. F., & Plass, C. (2001). Methylation matters. *Journal of Medical Genetics*, *38*(5), 285–303. https://doi.org/10.1136/jmg.38.5.285
Cretu, A., & Brooks, P. C. (2007). Impact of the non-cellular tumor microenvironment on metastasis: Potential therapeutic and imaging opportunities. *Journal of Cellular Physiology*, *213*(2), 391–402. https://doi.org/10.1002/jcp.21222
Dalton, R. S. J., Velineni, R., Osborne, M. E., Thomas, R., Harries, S., Gee, A. S., & Daniels, I. R. (2012). A single-centre experience of chemoradiotherapy for rectal cancer: Is there potential for nonoperative management? *Colorectal Disease*, *14*(5), 567–571. https://doi.org/10.1111/j.1463-1318.2011.02752.x
Davies, W., Isles, A. R., & Wilkinson, L. S. (2005). Imprinted gene expression in the brain. *Neuroscience and Biobehavioral Reviews*, *29*(3), 421–430. https://doi.org/10.1016/j.neubiorev.2004.11.007
Dbeis, R., Smart, N. J., & Daniels, I. R. (2016). Focusing the management of rectal cancer. *Annals of Translational Medicine*, *4*(24), 521. https://doi.org/10.21037/atm.2016.11.80
De Felice, F., Musio, D., Izzo, L., & Tombolini, V. (2014). Neoadjuvant chemoradiotherapy for locally advanced rectal cancer: The debate continues. *World Journal of Gastrointestinal Oncology*, *6*(12), 438–440. https://doi.org/10.4251/wjgo.v6.i12.438
De Maat, M. F. G., Van De Velde, C. J. H., Benard, A., Putter, H., Morreau, H., Van Krieken, J. H. J. M., … Hoon, D. S. B. (2010). Identification of a quantitative MINT locus methylation profile predicting local regional recurrence of rectal cancer. *Clinical Cancer Research*, *16*(10), 2811–2818. https://doi.org/10.1158/1078-0432.CCR-09-2717
De Maat, M. F. G., van De Velde, C. J. H., van Der Werff, M. P. J., Putter, H., Umetani, N., Klein-Kranenbarg, E. M., … Hoon, D. S. B. (2008). Quantitative analysis of methylation of genomic loci in early-stage rectal cancer predicts distant recurrence. *Journal of Clinical Oncology*, *26*(14), 2327–2335. https://doi.org/10.1200/JCO.2007.14.0723
de Ruijter, T. C., de Hoon, J. P. J., Slaats, J., de Vries, B., Janssen, M. J. F. W., van Wezel, T., … Veeck, J. (2015). Formalin-fixed, paraffin-embedded (FFPE) tissue epigenomics using Infinium HumanMethylation450 BeadChip assays. *Lab Invest*, *95*(7), 833–842. Retrieved from http://dx.doi.org/10.1038/labinvest.2015.53
Deaton, A. M., Webb, S., Kerr, A. R. W., Illingworth, R. S., Guy, J., Andrews, R., & Bird, A. (2011). Cell type–specific DNA methylation at intragenic CpG islands in the immune system. *Genome Research*, *21*(7), 1074–1086. https://doi.org/10.1101/gr.118703.110
Del Gobbo, A., & Ferrero, S. (2017). Immunohistochemical Markers as Predictors of Histopathologic Response and Prognosis in Rectal Cancer Treated with Preoperative Adjuvant Therapy: State of the Art. *Gastroenterology Research and Practice*, *2017*, 1–6. https://doi.org/10.1155/2017/2808235
Delmas, A. L., Riggs, B. M., Pardo, C. E., Dyer, L. M., Darst, R. P., & Izumchenko, E. G. (2011). WIF1 is a frequent target for epigenetic silencing in squamous cell carcinoma of the cervix. *Carcinogenesis*, *32*. https://doi.org/10.1093/carcin/bgr193
DeMorrow, S., Onori, P., Venter, J., Invernizzi, P., Frampton, G., White, M., & Franchitto, A. (2011). Neuropeptide Y inhibits cholangiocarcinoma cell

growth and invasion. *Am J Physiol Cell Physiol*, *300*. https://doi.org/10.1152/ajpcell.00358.2010

DeSantis, C. E., Lin, C. C., Mariotto, A. B., Siegel, R. L., Stein, K. D., Kramer, J. L., … Jemal, A. (2014). Cancer treatment and survivorship statistics, 2014. *CA: A Cancer Journal for Clinicians*, *64*(4), 252–271. https://doi.org/10.3322/caac.21235

Desch, C. E., Benson, A. B., Somerfield, M. R., Flynn, P. J., Krause, C., & Loprinzi, C. L. (2005). Colorectal cancer surveillance: 2005 update of an American society of clinical oncology practice guideline. *J Clin Oncol*, *23*. https://doi.org/10.1200/JCO.2005.04.0063

Devos, T., Tetzner, R., Model, F., Weiss, G., Schuster, M., Distler, J., … Lofton-Day, C. (2009). Circulating Methylated SEPT9 DNA in Plasma Is a Biomarker for Colorectal Cancer BT - Clin Chem.

Dewdney, A., Cunningham, D., Tabernero, J., Capdevila, J., Glimelius, B., Cervantes, A., … Chau, I. (2012). Multicenter Randomized Phase II Clinical Trial Comparing Neoadjuvant Oxaliplatin, Capecitabine, and Preoperative Radiotherapy With or Without Cetuximab Followed by Total Mesorectal Excision in Patients With High-Risk Rectal Cancer (EXPERT-C). *Journal of Clinical Oncology*, *30*(14), 1620–1627. https://doi.org/10.1200/JCO.2011.39.6036

Ding, Z., Qian, Y. B., Zhu, L. X., & Xiong, Q. R. (2009). Promoter methylation and mRNA expression of DKK-3 and WIF-1 in hepatocellular carcinoma. *World J Gastroenterol*, *15*. https://doi.org/10.3748/wjg.15.2595

Doerfler, W. (1983). DNA methylation and gene activity. *Ann. Rev. Biochem.*, *52*, 93–124. https://doi.org/10.1126/science.6254144

Doubeni, C. A., Major, J. M., Laiyemo, A. O., Schootman, M., Zauber, A. G., Hollenbeck, A. R., … Allison, J. (2012). Contribution of Behavioral Risk Factors and Obesity to Socioeconomic Differences in Colorectal Cancer Incidence. *JNCI Journal of the National Cancer Institute*, *104*(18), 1353–1362. https://doi.org/10.1093/jnci/djs346

Dulskas, A., Miliauskas, P., Tikuisis, R., Escalante, R., & Samalavicius, N. E. (2016). The functional results of radical rectal cancer surgery: Review of the literature. *Acta Chirurgica Belgica*, *116*(1), 1–10. https://doi.org/10.1080/00015458.2015.1136482

Ebert, M. P. A., Model, F., Mooney, S., Hale, K., Lograsso, J., Tonnes-Priddy, L., … Lofton-Day, C. (2006). Aristaless-like Homeobox-4 Gene Methylation Is a Potential Marker for Colorectal Adenocarcinomas. *Gastroenterology*, *131*(5), 1418–1430. https://doi.org/10.1053/j.gastro.2006.08.034

Ebert, M. P. A., Tänzer, M., Balluff, B., Burgermeister, E., Kretzschmar, A. K., Hughes, D. J., … Schmid, R. M. (2012). TFAP2E–DKK4 and Chemoresistance in Colorectal Cancer. *New England Journal of Medicine*, *366*(1), 44–53. https://doi.org/10.1056/NEJMoa1009473

Eckhardt, F., Lewin, J., Cortese, R., Rakyan, V. K., Attwood, J., Burger, M., … Beck, S. (2006). DNA methylation profiling of human chromosomes 6, 20 and 22, *38*(12), 1378–1385. Retrieved from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=17072317

Edge, S., Byrd, S., Compton, C., Et, A., & Editors. (2010). AJCC Cancer Staging Manual. In *7th edition* (pp. 143–164). New York, NY: Springer-Verlag.

Egger, G., Liang, G., Aparicio, A., & Jones, P. a. (2004). Epigenetics in human disease and prospects for epigenetic therapy. *Nature*, *429*(6990), 457–63. https://doi.org/10.1038/nature02625

Ehrlich, M. (2002). DNA methylation in cancer: too much, but also too little. *Oncogene*, *21*(35), 5400–5413. https://doi.org/10.1038/sj.onc.1205651

Eide, T., Ramberg, H., Glackin, C., Tindall, D., & Taskén, K. A. (2013). TWIST1, A novel androgen-regulated gene, is a target for NKX3-1 in prostate cancer cells. *Cancer Cell International*, *13*(1), 4. https://doi.org/10.1186/1475-2867-13-4

Enblad, P., Adami, H.-O., Bergström, R., Glimelius, B., Krusemo, U., & Påhlman, L. (1988). Improved Survival of Patients With Cancers of the Colon and Rectum?2. *JNCI: Journal of the National Cancer Institute*, *80*(8), 586–591. Retrieved from http://dx.doi.org/10.1093/jnci/80.8.586

ENCODEProjectConsortium, Dunham, I., Kundaje, A., Aldred, S. F., Collins, P. J., Davis, C. A., … Lochovsky, L. (2012). An integrated encyclopedia of {DNA} elements in the human genome. *Nature*, *489*(7414), 57–74. https://doi.org/10.1038/nature11247

Esteller, M. (2005). DNA methylation and cancer therapy: new developments and expectations. *Curr Opin Oncol*, *17*. https://doi.org/10.1097/01.cco.0000147383.04709.10

Etcheverry, A., Aubry, M., de Tayrac, M., Vauleon, E., Boniface, R., Guenot, F., … Mosser, J. (2010). DNA methylation in glioblastoma: impact on gene expression and clinical outcome. *BMC Genomics*, *11*, 701. https://doi.org/10.1186/1471-2164-11-701

European Bioinformatics Institute (2018). GO:0031012 extracellular matrix. *In (EMBL-EBI)*. Retrieved from https://www.ebi.ac.uk/QuickGO/term/GO:0005578

Exner, R., Pulverer, W., Diem, M., Spaller, L., Woltering, L., Schreiber, M., … Egger, G. (2015). Potential of DNA methylation in rectal cancer as diagnostic and prognostic biomarkers. *British Journal of Cancer*, *113*(7), 1035–1045. https://doi.org/10.1038/bjc.2015.303

Fackler, M. J., Malone, K., Zhang, Z., Schilling, E., Garrett-Mayer, E., & Swift-Scanlan, T. (2006). Quantitative multiplex methylation-specific PCR analysis doubles detection of tumor cells in breast ductal fluid. *Clin Cancer Res*, *12*. https://doi.org/10.1158/1078-0432.CCR-05-2733

Faivre, J., Dancourt, V., Lejeune, C., Tazi, M. A., Lamour, J., & Gerard, D. (2004). Reduction in colorectal cancer mortality by fecal occult blood screening in a French controlled study. *Gastroenterology*, *126*. https://doi.org/10.1053/j.gastro.2004.02.018

Fazeli, M. S., & Keramati, M. R. (2015). Rectal cancer: a review. *Medical Journal of the Islamic Republic of Iran*, *29*, 171. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4431429/

Fearon, E. R. (2011). Molecular genetics of colorectal cancer. *Annual Review of Pathology*, *6*, 479–507. https://doi.org/10.1146/annurev-pathol-011110-130235

Feinberg, A. P., & Tycko, B. (2004). The history of cancer epigenetics. *Nature Reviews. Cancer*, *4*(2), 143–53. https://doi.org/10.1038/nrc1279

Feinberg, A. P., & Vogelstein, B. (1983). Hypomethylation distinguishes genes of some human cancers from their normal counterparts. *Nature*, *301*(5895), 89–92. https://doi.org/10.1038/301089a0

Feinberg, A. P. (2004). The epigenetics of cancer etiology. *Seminars in Cancer Biology*, *14*(6), 427–432. https://doi.org/10.1016/j.semcancer.2004.06.005

Fend, F., & V, M. R. (2000). Laser capture microdissection in pathology, 666–672.
Figueredo, A., Coombes, M. E., & Mukherjee, S. (2008). Adjuvant Therapy for completely resected Stage II Colon Cancer. *Cochrane Database of Systematic Reviews*, (3). Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=c8h&AN=2009993146&site=ehost-live
Figueredo, A., Coombes, M. E., & Mukherjee, S. (2008). Adjuvant Therapy for completely resected Stage II Colon Cancer. *Cochrane Database of Systematic Reviews*, (3). https://doi.org/10.1002/14651858.CD005390.pub2
Finlay, J., Roberts, C. M., Lowe, G., Loeza, J., Rossi, J. J., & Glackin, C. A. (2015). RNA-based TWIST1 inhibition via dendrimer complex to reduce breast cancer cell metastasis. *BioMed Research International*, *2015*. https://doi.org/10.1155/2015/382745
Fodde, R., Smits, R., & Clevers, H. (2001). Signal transduction and genetic instability in colorectal cancer. *October*, *1*(October).
Fodde, R., Smits, R., & Clevers, H. (2001). APC, Signal transduction and genetic instability in colorectal cancer. *Nat Rev Cancer*, *1*(1), 55–67. Retrieved from http://dx.doi.org/10.1038/35094067
Fodde, R., & Tomlinson, I. (2010). Nuclear β-catenin expression and Wnt signalling: in defence of the dogma. *The Journal of Pathology*, *221*(3), 239–241. https://doi.org/10.1002/path.2718
Folkesson, J., Birgisson, H., Pahlman, L., Cedermark, B., Glimelius, B., & Gunnarsson, U. (2005). Swedish Rectal Cancer Trial: Long Lasting Benefits From Radiotherapy on Survival and Local Recurrence Rate. *Journal of Clinical Oncology*, *23*(24), 5644–5650. https://doi.org/10.1200/JCO.2005.08.144
FOxTROT Collaborative, G. (2017). Feasibility of preoperative chemotherapy for locally advanced, operable colon cancer: the pilot phase of a randomised controlled trial. *The Lancet Oncology*, *13*(11), 1152–1160. https://doi.org/10.1016/S1470-2045(12)70348-0
Frattini, M., Balestra, D., Suardi, S., Oggionni, M., Alberici, P., Radice, P., … Pierotti, M. A. (2004). Different genetic features associated with colon and rectal carcinogenesis. *Clinical Cancer Research*, *10*(12 I), 4015–4021. https://doi.org/10.1158/1078-0432.CCR-04-0031
Frazier, A. L., Colditz, G. A., Fuchs, C. S., & Kuntz, K. M. (2000). Cost-effectiveness of screening for colorectal cancer in the general population. *JAMA*, *284*. https://doi.org/10.1001/jama.284.15.1954
Friedrich, M. J., Rad, R., Langer, R., Voland, P., Hoefler, H., Schmid, R. M., … Gerhard, M. (2006). Lack of RUNX3 regulation in human gastric cancer. *The Journal of Pathology*, *210*(2), 141–146. https://doi.org/10.1002/path.2042
Frigola, J., Song, J., Stirzaker, C., Hinshelwood, R. A., Peinado, M. A., & Clark, S. J. (2006). Epigenetic remodeling in colorectal cancer results in coordinate gene suppression across an entire chromosome band. *Nature Genetics*, *38*(5), 540–549. https://doi.org/10.1038/ng1781
Frommer, M., McDonald, L. E., Millar, D. S., Collis, C. M., Watt, F., Grigg, G. W., … Paul, C. L. (1992). A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. *Proceedings of the National Academy of Sciences*, *89*(5), 1827–1831. https://doi.org/10.1073/pnas.89.5.1827

Fukushima, N., Sato, N., & Ueki, T. (2002). Aberrant methylation of preproenkephalin and p16 genes in pancreatic intraepithelial neoplasia and pancreatic ductal adenocarcinoma. *Am J Pathol*, *160*. https://doi.org/10.1016/S0002-9440(10)61104-2

Gaedcke, J., Leha, A., Claus, R., Weichenhan, D., Jung, K., Kitz, J., … Ghadimi, M. (2014). Identification of a DNA methylation signature to predict disease-free survival in locally advanced rectal cancer. *Oncotarget*, *5*(18), 8123–35.

Gardiner-Garden, M., & Frommer, M. (1987). CpG Islands in vertebrate genomes. *Journal of Molecular Biology*, *196*(2), 261–282. https://doi.org/http://dx.doi.org/10.1016/0022-2836(87)90689-9

Gerecke, C., Scholtka, B., Löwenstein, Y., Fait, I., Gottschalk, U., Rogoll, D., … Kleuser, B. (2015). Hypermethylation of ITGA4, TFPI2 and VIMENTIN promoters is increased in inflamed colon tissue: putative risk markers for colitis-associated cancer. *Journal of Cancer Research and Clinical Oncology*, *141*(12), 2097–2107. https://doi.org/10.1007/s00432-015-1972-8

Giotopoulos, G., McCormick, C., Cole, C., Zanker, A., Jawad, M., Brown, R., & Plumb, M. (2006). DNA methylation during mouse hemopoietic differentiation and radiation-induced leukemia. *Experimental Hematology*, *34*(11), 1462–1470. https://doi.org/10.1016/j.exphem.2006.06.008

Glimelius, B. (2012). Which Rectal Cancers Are Locally Advanced? | Cancer Network. *Oncology*. Retrieved from http://www.cancernetwork.com/colorectal-cancer/which-rectal-cancers-are-locally-advanced

Glimelius, B., Tiret, E., Cervantes, A., & Arnold, D. (2013). Rectal cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up†. *Annals of Oncology*, *24*(suppl_6), vi81-vi88. Retrieved from http://dx.doi.org/10.1093/annonc/mdt240

Glockner, S. C., Dhir, M., Yi, J. M., McGarvey, K. E., Van Neste, L., Louwagie, J., … Ahuja, N. (2009). Methylation of TFPI2 in Stool DNA: A Potential Novel Biomarker for the Detection of Colorectal Cancer. *Cancer Research*, *69*(11), 4691–4699. https://doi.org/10.1158/0008-5472.CAN-08-0142

Goel, A., & Boland, C. R. (2012). Epigenetics of Colorectal Cancer. *Gastroenterology*, *143*(6), 1442–1460.e1. https://doi.org/10.1053/j.gastro.2012.09.032

Goo, Y. A., Goodlett, D. R., & Pascal, L. E. (2005). Stromal mesenchyme cell genes of the human prostate and bladder. *BMC Urol*, *5*. https://doi.org/10.1186/1471-2490-5-17

Gozzi, G., Chelbi, S. T., Manni, P., Alberti, L., Fonda, S., Saponaro, S., … Losi, L. (2016). Promoter methylation and downregulated expression of the TBX15 gene in ovarian carcinoma. *Oncology Letters*, *12*(4), 2811–2819. https://doi.org/10.3892/ol.2016.5019

GRADY, W. M., & CARETHERS, J. M. (2008). Genomic and Epigenetic Instability in Colorectal Cancer Pathogenesis. *Gastroenterology*, *135*(4), 1079–1099. https://doi.org/10.1053/j.gastro.2008.07.076

Gregorieff, A., Pinto, D., Begthel, H., Destree, O., Kielman, M., & Clevers, H. (2005). Expression pattern of Wnt signaling components in the adult intestine. *Gastroenterology*, *129*. https://doi.org/10.1016/j.gastro.2005.06.007

Grützmann, R., Molnar, B., Pilarsky, C., Habermann, J. K., Schlag, P. M., Saeger, H. D., … Lofton-Day, C. (2008). Sensitive detection of colorectal

cancer in peripheral blood by septin 9 DNA methylation assay. *PLoS ONE*, *3*(11). https://doi.org/10.1371/journal.pone.0003759
Gupta, S., Ashfaq, R., Kapur, P., Afonso, B. B., Nguyen, T. P., Ansari, F., … Rockey, D. C. (2011). Microsatellite instability among individuals of Hispanic origin with colorectal cancer. *Cancer*, *116*. https://doi.org/10.1002/cncr.25486
Gurpinar, E., Grizzle, W. E., & Piazza, G. A. (2014). NSAIDs inhibit tumorigenesis, but how? *Clinical Cancer Research : An Official Journal of the American Association for Cancer Research*, *20*(5), 1104–1113. https://doi.org/10.1158/1078-0432.CCR-13-1573
Yang, Y.-X. (2012). Identification of novel DNA methylation markers in colorectal cancer using MIRA-based microarrays. *Oncology Reports*, *28*(1), 99–104. https://doi.org/10.3892/or.2012.1779
Habano, W., Gamo, T., Sugai, T., Otsuka, K., Wakabayashi, G., & Ozawa, S. (2009). CYP1B1, but not CYP1A1, is downregulated by promoter methylation in colorectal cancers. *International Journal of Oncology*, *34*(4), 1085–1091. https://doi.org/10.3892/ijo_00000235
Habr-Gama, A., Perez, R. O., Nadalin, W., Sabbaga, J., Ribeiro, U., Silva e Sousa, A. H., … Gama-Rodrigues, J. (2004). Operative Versus Nonoperative Treatment for Stage 0 Distal Rectal Cancer Following Chemoradiation Therapy: Long-term Results. *Annals of Surgery*, *240*(4), 711–718. https://doi.org/10.1097/01.sla.0000141194.27992.32
Hahn, M. A., Hahn, T., Lee, D. H., Esworthy, R. S., Kim, B. W., Riggs, A. D., … Pfeifer, G. P. (2008). Methylation of polycomb target genes in intestinal cancer is mediated by inflammation. *Cancer Research*, *68*(24), 10280–10289. https://doi.org/10.1158/0008-5472.CAN-08-1957
Hallden, G., Hadi, M., Hong, H. T., & Aponte, G. W. (1999). Y receptor-mediated induction of CD63 transcripts, a tetraspanin determined to be necessary for differentiation of the intestinal epithelial cell line, hBRIE 380i cells. *J Biol Chem*, *274*. https://doi.org/10.1074/jbc.274.39.27914
Hamilton, W., Coleman, M. G., & Rubin, G. (2013). Colorectal cancer. *BMJ : British Medical Journal*, *346*. Retrieved from http://www.bmj.com/content/346/bmj.f3172.abstract
Hammoud, S. S., Cairns, B. R., & Jones, D. A. (2013). Epigenetic regulation of colon cancer and intestinal stem cells. *Current Opinion in Cell Biology*, *25*(2), 177–183. https://doi.org/10.1016/j.ceb.2013.01.007
Han, Z., Liu, L., Liu, Y., & Li, S. (2014). Sirtuin SIRT6 suppresses cell proliferation through inhibition of Twist1 expression in non-small cell lung cancer. *International Journal of Clinical and Experimental Pathology*, *7*(8), 4774–4781.
Hardcastle, J. D., Thomas, W. M., Chamberlain, J., Pye, G., Sheffield, J., & James, P. D. (1989). Ramdomised, controlled trial of faecal occult blood screening for colorectal cancer: results for first 107,349 subjects. *Lancet*, *1*. https://doi.org/10.1016/S0140-6736(89)92750-5
Hardwick, J. C., Kodach, L. L., Offerhaus, G. J., & van Den Brink, G. R. (2008). Bone morphogenetic protein signalling in colorectal cancer. *Nature Reviews. Cancer*, *8*, 806–812.
Hardwick, J., Kodach, L., Offerhaus, G., & Van Den Brink, G. (2004). Bone Morphogenetic protein signalling in colorectal cancer. *Nature Reviews Cancer*, *1*, 111–21.

Harrison, S., & Benziger, H. (2017). The molecular biology of colorectal carcinoma and its implications: A review. *The Surgeon*, *9*(4), 200–210. https://doi.org/10.1016/j.surge.2011.01.011
Harriss, D. J., Atkinson, G., Batterham, A., George, K., Tim Cable, N., Reilly, T., … The Colorectal Cancer Exercise And Research Group, L. (2009). Lifestyle factors and colorectal cancer risk (2): a systematic review and meta-analysis of associations with leisure-time physical activity. *Colorectal Disease*, *11*(7), 689–701. https://doi.org/10.1111/j.1463-1318.2009.01767.x
He, B., Reguart, N., You, L., Mazieres, J., Xu, Z., & Lee, A. Y. (2005). Blockade of Wnt-1 signaling induces apoptosis in human colorectal cancer cells containing downstream mutations. *Oncogene*, *24*. https://doi.org/10.1038/sj.onc.1208511
He, Q., Chen, H. Y., Bai, E. Q., Luo, Y. X., Fu, R. J., He, Y. S., … Wang, H. Q. (2010). Development of a multiplex MethyLight assay for the detection of multigene methylation in human colorectal cancer. *Cancer Genet Cytogenet*, *202*. https://doi.org/10.1016/j.cancergencyto.2010.05.018
HE, S.-Y., JIANG, R.-F., JIANG, J. I. E., XIANG, Y.-S., & WANG, L. (2015). Investigation of methylation and protein expression of the Runx3 gene in colon carcinogenesis. *Biomedical Reports*, *3*(5), 687–690. https://doi.org/10.3892/br.2015.479
Hellebrekers, D. M., Melotte, V., Vire, E., Langenkamp, E., Molema, G., Fuks, F., … van Engeland, M. (2007). Identification of epigenetically silenced genes in tumor endothelial cells. *Cancer Res*, *67*. https://doi.org/10.1158/0008-5472.CAN-06-3032
Hendrich, B., & Bird, A. (1998). Identification and characterization of a family of mammalian methyl-CpG binding proteins. *Molecular and Cellular Biology*, *18*(11), 6538–47. https://doi.org/10.1128/MCB.18.11.6538
Henikoff, S., & Matzke, M. A. (1997). Exploring and explaining epigenetic effects. *Trends in Genetics : TIG*, *13*(8), 293–295. https://doi.org/10.1016/S0168-9525(97)01219-5
Herbst, A., Rahmig, K., Stieber, P., Philipp, A., Jung, A., Ofner, A., … Kolligs, F. T. (2011). Methylation of NEUROG1 in Serum Is a Sensitive Marker for the Detection of Early Colorectal Cancer. *Am J Gastroenterol*, *106*. https://doi.org/10.1038/ajg.2011.6
Hibi, K., Goto, T., Shirahata, A., Saito, M., Kigawa, G., Nemoto, H., & Sanada, Y. (2012). Methylation of TFPI2 no longer detected in the serum DNA of colorectal cancer patients after curative surgery. *Anticancer Research*, *32*(3), 787–790.
Hibi, K., Mizukami, H., Saito, M., Kigawa, G., Nemoto, H., & Sanada, Y. (2012). FBN2 methylation is detected in the serum of colorectal cancer patients with hepatic metastasis. *Anticancer Research*, *32*(10), 4371–4374.
Hinoue, T., Weisenberger, D. J., Lange, C. P. E., Shen, H., Byun, H.-M., Van Den Berg, D., … Laird, P. W. (2012). Genome-scale analysis of aberrant DNA methylation in colorectal cancer. *Genome Research*, *22*(2), 271–282. https://doi.org/10.1101/gr.117523.110
HMGov. (2004). Human Tissue Act. *HMSO*. HMSO. https://doi.org/10.1136/bmj.1.5590.516-a
Hoeller, D., Hecker, C.-M., & Dikic, I. (2006). Ubiquitin and ubiquitin-like proteins in cancer pathogenesis. *Nat Rev Cancer*, *6*(10), 776–788. Retrieved from http://dx.doi.org/10.1038/nrc1994
Hofstetter, B., Niemierko, A., Forrer, C., Benhattar, J., Albertini, V., Pruschy, M., … Ciernik, I. F. (2010). Impact of Genomic Methylation on Radiation

Sensitivity of Colorectal Carcinoma. *International Journal of Radiation Oncology Biology Physics*, *76*(5), 1512–1519. https://doi.org/10.1016/j.ijrobp.2009.10.037
Hohenberger, W., Weber, K., Matzel, K., Papadopoulos, T., & Merkel, S. (2009). Standardized surgery for colonic cancer: complete mesocolic excision and central ligation – technical notes and outcome. *Colorectal Disease*, *11*(4), 354–364. https://doi.org/10.1111/j.1463-1318.2008.01735.x
Hoque, M. O., Begum, S., Topaloglu, O., Chatterjee, A., Rosenbaum, E., & Van, C. W. (2006). Quantitation of promoter methylation of multiple genes in urine DNA and bladder cancer detection. *J Natl Cancer Inst*, *98*. https://doi.org/10.1093/jnci/djj265
Horvath, S., Cabras, S., Marini, E., Lauber, C., Knights, D., Ackermann, G., … Berg-Lyons, D. (2013). DNA methylation age of human tissues and cell types. *Genome Biology*, *14*(10), R115. https://doi.org/10.1186/gb-2013-14-10-r115
Howard, G., Eiges, R., Gaudet, F., Jaenisch, R., & Eden, A. (2008). Activation and transposition of endogenous retroviral elements in hypomethylation induced tumors in mice. *Oncogene*, *27*. https://doi.org/10.1038/sj.onc.1210631
Hsieh, J. C., Kodjabachian, L., Rebbert, M. L., Rattner, A., Smallwood, P. M., & Samos, C. H. (1999). A new secreted protein that binds to Wnt proteins and inhibits their activities. *Nature*, *398*. https://doi.org/10.1038/18899
Hua, Y., Ma, X., Liu, X., Yuan, X., Qin, H., & Zhang, X. (2017). Abnormal expression of mRNA, microRNA alteration and aberrant DNA methylation patterns in rectal adenocarcinoma. *PLoS ONE*, *12*(3), e0174461. https://doi.org/10.1371/journal.pone.0174461
Huang, T., Kang, W., Cheng, A. S. L., Yu, J., & To, K. F. (2015). The emerging role of Slit-Robo pathway in gastric and other gastro intestinal cancers. *BMC Cancer*, *15*, 950. https://doi.org/10.1186/s12885-015-1984-4
Huang, W. Y., Hsu, S. D., Huang, H. Y., Sun, Y. M., Chou, C. H., Weng, S. L., & Huang, H. D. (n.d.). MethHC: a database of DNA methylation and gene expression in human cancer. *Nucleic Acids Res*, *43*(Database issue), D856-61. Retrieved from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=25398901
Huang, X., Xiao, D. W., Xu, L. Y., Zhong, H. J., Liao, L. D., Xie, Z. F., & Li, E. M. (2009). Prognostic significance of altered expression of SDC2 and CYR61 in esophageal squamous cell carcinoma. *Oncol Rep*, *21*(4), 1123–1129. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/19288017
Huang, Z., Li, L., & Wang, J. (2007). Hypermethylation of SFRP2 as a potential marker for stool-based detection of colorectal cancer and precancerous lesions. *Dig Dis Sci*, *52*. https://doi.org/10.1007/s10620-007-9755-y
Hynes, R. O. (2009). The Extracellular Matrix: Not Just Pretty Fibrils. *Science*, *326*(5957), 1216–1219. https://doi.org/10.1126/science.1176009
Iacopetta, B. (2002). Are there two sides to colorectal cancer? *International Journal of Cancer*, *101*(5), 403–408. https://doi.org/10.1002/ijc.10635
IARC, & Cancer, I. A. for R. on. (2014). *World Cancer Report 2014*. Lyon, France.
Ibrahim, A. E., Arends, M. J., Silva, A. L., Wyllie, A. H., Greger, L., Ito, Y., … Brenton, J. D. (2011). Sequential DNA methylation changes are associated with DNMT3B overexpression in colorectal neoplastic progression. *Gut*, *60*. https://doi.org/10.1136/gut.2010.223602

Irizarry, R. A., Bravo, H. C., Irizarry, R. A., Irizarry, R., Hobbs, B., Collin, F., … Durbin, R. (2003). Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics*, *4*(2), 249–264. https://doi.org/10.1093/biostatistics/4.2.249

Irizarry, R. A., Ladd-Acosta, C., Wen, B., Wu, Z., Montano, C., Onyango, P., … Feinberg, A. P. (2009). Genome-wide methylation analysis of human colon cancer reveals similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. *Nature Genetics*, *41*(2), 178–186. https://doi.org/10.1038/ng.298

Irving, M. H., & Catchpole, B. (1992). ABC of colorectal diseases. Anatomy and physiology of the colon, rectum, and anus. *BMJ : British Medical Journal*, *304*(6834), 1106–1108. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1881899/

Itzkowitz, S., Brand, R., Jandorf, L., Durkee, K., Millholland, J., Rabeneck, L., … Berger, B. M. (2008). A simplified, noninvasive Stool DNA test for colorectal cancer detection. *American Journal of Gastroenterology*, *103*(11), 2862–2870. https://doi.org/10.1111/j.1572-0241.2008.02088.x

Iwamoto, K. S., Mizuno, T., Ito, T., Akiyama, M., Takeichi, N., Mabuchi, K., & Seyama, T. (1996). Feasibility of using decades-old archival tissues in molecular oncology/epidemiology. *The American Journal of Pathology*, *149*(2), 399–406.

Jaenisch, R., & Bird, A. (2003). Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nature Genetics*, *33*(3s), 245–254. https://doi.org/10.1038/ng1089

Jemal, A., Siegel, R., Ward, E., Hao, Y., Xu, J., Murray, T., & Thun, M. J. (2008). Cancer statistics, 2008. *CA Cancer J Clin*, *58*. https://doi.org/10.3322/CA.2007.0010

Jemal, A., Siegel, R., Xu, J., & Ward, E. (2010). Cancer Statistics, 2010. *Cancer J Clin*, *60*. https://doi.org/10.3322/caac.20073

Jo, P., Jung, K., Grade, M., Conradi, L. C., Wolff, H. A., Kitz, J., … Gaedcke, J. (2012). CpG island methylator phenotype infers a poor disease-free survival in locally advanced rectal cancer. *Surgery*, *151*(4), 564–570. https://doi.org/10.1016/j.surg.2011.08.013

Jo, P., Jung, K., Grade, M., Conradi, L.-C., Wolff, H. A., Kitz, J., … Gaedcke, J. (2017). CpG island methylator phenotype infers a poor disease-free survival in locally advanced rectal cancer. *Surgery*, *151*(4), 564–570. https://doi.org/10.1016/j.surg.2011.08.013

Jones, P. A. (2002). DNA methylation and cancer. *Oncogene*, *21*. https://doi.org/10.1038/sj.onc.1205597

Jorge, J., & Habr-Gama, A. (2014). Anatomy and Embryology of the Colon, Rectum, and Anus BT - The ASCRS Manual of Colon and Rectal Surgery.

Jorge, J. M. N., & Habr-Gama, A. (2014). Anatomy and Embryology of the Colon, Rectum, and Anus BT - The ASCRS Manual of Colon and Rectal Surgery. In D. E. Beck, S. D. Wexner, T. L. Hull, P. L. Roberts, T. J. Saclarides, A. J. Senagore, … S. R. Steele (Eds.) (pp. 1–25). New York, NY: Springer New York. https://doi.org/10.1007/978-1-4614-8450-9_1

Jowaed, A., Schmitt, I., Kaut, O., & Wullner, U. (2010). Methylation regulates alpha-synuclein expression and is decreased in Parkinson's disease patients' brains. *J Neurosci*, *30*. https://doi.org/10.1523/JNEUROSCI.6119-09.2010

Jung, A. (2013). Colon and Rectum Normal Anatomy and Congenital Variants. In B. Hamm & P. R. Ros (Eds.), *Abdominal Imaging* (pp. 785–796). Berlin,

Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-13327-5_155
Kachnic, L. A. (2007). Adjuvant Chemoradiation for Localized Rectal Cancer: Current Trends and Future Directions. *Gastrointestinal Cancer Research : GCR*, *1*(4 Suppl 2), S64–S72. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2666832/
Kalra, S. P., & Kalra, P. S. (2003). Neuropeptide Y: a physiological orexigen modulated by the feedback action of ghrelin and leptin. *Endocrine*, *22*. https://doi.org/10.1385/ENDO:22:1:49
Kamiyama, H., Suzuki, K., Maeda, T., Koizumi, K., Miyaki, Y., Okada, S., … Perucho, M. (2012). DNA demethylation in normal colon tissue predicts predisposition to multiple cancers. *Oncogene*, *31*. https://doi.org/10.1038/onc.2011.652
Kanellos, D., Kitsios, G., Kanellos, I., Demetriades, H., Pramateftakis, M. G., Angelopoulos, S., & Betsis, D. (2004). Anaemia as a symptom of right colon cancer. *Techniques in Coloproctology*, *8*(1), s62–s64. https://doi.org/10.1007/s10151-004-0114-0
Kang, S. K., Chae, Y. K., Woo, J., Kim, M. S., Park, J. C., Lee, J., … Moon, C. (2008). Role of Human Aquaporin 5 In Colorectal Carcinogenesis. *The American Journal of Pathology*, *173*(2), 518–525. https://doi.org/10.2353/ajpath.2008.071198
Kansara, M., Tsang, M., Kodjabachian, L., Sims, N. A., Trivett, M. K., & Ehrich, M. (2009). Wnt inhibitory factor 1 is epigenetically silenced in human osteosarcoma, and targeted disruption accelerates osteosarcomagenesis in mice. *Clin Invest*, *119*. https://doi.org/10.1172/JCI37175
Kapiteijn, E., Liefers, G. J., Los, L. C., Klein Kranenbarg, E., Hermans, J., Tollenaar, R. A. E. M., … van Krieken, J. H. J. M. (2001). Mechanisms of oncogenesis in colon versus rectal cancer. *The Journal of Pathology*, *195*(2), 171–178. https://doi.org/10.1002/path.918
Kapranov, P., St Laurent, G., Raz, T., Ozsolak, F., Reynolds, C. P., Sorensen, P. H., … Triche, T. J. (2010). The majority of total nuclear-encoded non-ribosomal RNA in a human cell is “dark matter” un-annotated RNA. *BMC Biology*, *8*(1), 149. https://doi.org/10.1186/1741-7007-8-149
Katz, M. S., Minsky, B. D., Saltz, L. B., Riedel, E., Chessin, D. B., & Guillem, J. G. (2017). Association of statin use with a pathologic complete response to neoadjuvant chemoradiation for rectal cancer. *International Journal of Radiation Oncology • Biology • Physics*, *62*(5), 1363–1370. https://doi.org/10.1016/j.ijrobp.2004.12.033
Katz, M. S., Minsky, B. D., Saltz, L. B., Riedel, E., Chessin, D. B., & Guillem, J. G. (2005). Association of statin use with a pathologic complete response to neoadjuvant chemoradiation for rectal cancer. *International Journal of Radiation Oncology Biology Physics*, *62*(5), 1363–1370. https://doi.org/10.1016/j.ijrobp.2004.12.033
Kawamoto, K., Onodera, H., Kondo, S., Kan, S., Ikeuchi, D., Maetani, S., & Imamura, M. (1998). Expression of insulin-like growth factor-2 can predict the prognosis of human colorectal cancer patients: correlation with tumor progression, proliferative activity and survival. *Oncology*, *55*(3), 242–248. https://doi.org/ocl55242 [pii]
Kazazian, H. H. (2004). Mobile elements: drivers of genome evolution. *Science*, *303*. https://doi.org/10.1126/science.1089670
Kealey, S. M., Dodd, J. D., MacEneaney, P. M., Gibney, R. G., & Malone, D. E. (2017). Minimal preparation computed tomography instead of barium

enema/colonoscopy for suspected colon cancer in frail elderly patients: an outcome analysis study. *Clinical Radiology*, *59*(1), 44–52. https://doi.org/10.1016/j.crad.2003.08.001

Kent, W. J., Sugnet, C. W., Furey, T. S., Roskin, K. M., Pringle, T. H., Zahler, A. M., & Haussler, D. (2002). The human genome browser at UCSC. *Genome Res*, *12*. https://doi.org/10.1101/gr.229102. Article published online before print in May 2002

Kilpinen, S., Autio, R., Ojala, K., Iljin, K., Bucher, E., Sara, H., … Kallioniemi, O. (2008). Systematic bioinformatic analysis of expression levels of 17,330 human genes across 9,783 samples from 175 types of healthy and pathological tissues. *Genome Biol*, *9*. https://doi.org/10.1186/gb-2008-9-9-r139

Kim, B. R., Kang, M. H., Kim, J. L., Na, Y. J., Park, S. H., Lee, S. Il, … Oh, S. C. (2016). RUNX3 inhibits the metastasis and angiogenesis of colorectal cancer. *Oncology Reports*, *36*(5), 2601–2608. https://doi.org/10.3892/or.2016.5086

Kim, M. S., Chang, X., Nagpal, J. K., Yamashita, K., Baek, J. H., Dasgupta, S., … Sidransky, D. (2008). The N-methyl-D-aspartate receptor type 2A is frequently methylated in human colorectal carcinoma and suppresses cell growth. *Oncogene*, *27*. https://doi.org/10.1038/sj.onc.1210842

Kim, M. S., Louwagie, J., Carvalho, B., Terhaar Sive Droste, J. S., Park, H. L., Chae, Y. K., … Sidransky, D. (2009). Promoter DNA methylation of oncostatin m receptor-beta as a novel diagnostic and therapeutic marker in colon cancer. *PLoS ONE*, *4*. https://doi.org/10.1371/journal.pone.0006555

Kim, M. S., Lee, J., & Sidransky, D. (2010). DNA methylation markers in colorectal cancer. *Cancer and Metastasis Reviews*, *29*(1), 181–206. https://doi.org/10.1007/s10555-010-9207-6

Kim, S. H., Chang, H. J., Kim, D. Y., Park, J. W., Baek, J. Y., Kim, S. Y., … Nam, B.-H. (2016). What is the ideal tumor regression grading system in rectal cancer patients after preoperative chemoradiotherapy? *Cancer Research and Treatment*, *48*(3). https://doi.org/10.4143/crt.2015.254

Kim, S. J., Tae, C. H., Hong, S. N., Min, B. H., Chang, D. K., Rhee, P. L., … Kim, Y. H. (2015). EYA4 Acts as a New Tumor Suppressor Gene in Colorectal Cancer. *Molecular Carcinogenesis*, *54*(12), 1748–1757. https://doi.org/10.1002/mc.22247

Kim, W., Vo, Q., Shrivastav, M., Lataxes, T., & Brown, K. (2002). Aberrant methylation of the ATM promoter correlates with increased radiosensitivity in a human colorectal tumor cell line. *Oncogene*, (October 2001), 3864–3871. https://doi.org/10.1038/sj/onc/1205485

Kim, Y. H., Petko, Z., Dzieciatkowski, S., Lin, L., Ghiassi, M., Stain, S., … Grady, W. M. (2006). CpG island methylation of genes accumulates during the adenoma progression step of the multistep pathogenesis of colorectal cancer. *Genes Chromosomes and Cancer*, *45*(8), 781–789. https://doi.org/10.1002/gcc.20341

Kim, Y.-H., Lee, H. C., Kim, S.-Y., Yeom, Y. Il, Ryu, K. J., Min, B.-H., … Kim, Y. S. (2011). Epigenomic Analysis of Aberrantly Methylated Genes in Colorectal Cancer Identifies Genes Commonly Affected by Epigenetic Alterations. *Annals of Surgical Oncology*, *18*(8), 2338–2347. https://doi.org/10.1245/s10434-011-1573-y

Kleivi, K., Teixeira, M. R., Eknaes, M., Diep, C. B., Jakobsen, K. S., Hamelin, R., & Lothe, R. A. (2004). Genome signatures of colon carcinoma cell lines.

Cancer Genet Cytogenet, *155*. https://doi.org/10.1016/j.cancergencyto.2004.03.014

Kohonen-Corish, M. R. J., Tseung, J., Chan, C., Currey, N., Dent, O. F., Clarke, S., … Chapuis, P. H. (2014). KRAS mutations and CDKN2A promoter methylation show an interactive adverse effect on survival and predict recurrence of rectal cancer. *International Journal of Cancer*, *134*(12), 2820–2828. https://doi.org/10.1002/ijc.28619

Kondraganti, S., Gondi, C. S., Gujrati, M., McCutcheon, I., Dinh, D. H., Rao, J. S., & Olivero, W. C. (2006). Restoration of tissue factor pathway inhibitor (TFPI) inhibits invasion and tumor growth in vitro and in vivo in a malignant meningioma cell line. *International Journal of Oncology*, *29*(1), 25–32. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1479607/

Kosinski, L., Habr-Gama, A., Ludwig, K., & Perez, R. (2012). Shifting concepts in rectal cancer management. *CA: A Cancer Journal for Clinicians*, *62*(3), 173–202. https://doi.org/10.3322/caac.21138

Kouzarides, T., Aggarwal, B. D., Calvi, B. R., Ahn, S. H., Cheung, W. L., Hsu, J. Y., … Reinberg, D. (2007). Chromatin modifications and their function. *Cell*, *128*(4), 693–705. https://doi.org/10.1016/j.cell.2007.02.005

Kuepper, C., Großerueschkamp, F., Kallenbach-Thieltges, A., Mosig, A., Tannapfel, A., & Gerwert, K. (2016). Label-free classification of colon cancer grading using infrared spectral histopathology. *Faraday Discuss.*, *187*, 105–118. https://doi.org/10.1039/C5FD00157A

Kuipers, E. J., Grady, W. M., Lieberman, D., Seufferlein, T., Sung, J. J., Boelens, P. G., … Watanabe, T. (2015). Colorectal cancer. *Nature Reviews Disease Primers*, *1*, 15065. Retrieved from http://dx.doi.org/10.1038/nrdp.2015.65

Kunkel, L. M., Smith, K. D., Boyer, S. H., Borgaonkar, D. S., Wachtel, S. S., Miller, O. J., … Rary, J. M. (1977). Analysis of human Y-chromosome-specific reiterated DNA in chromosome variants. *Proc Natl Acad Sci USA*, *74*. https://doi.org/10.1073/pnas.74.3.1245

Kuo, C.-C., Shih, Y.-L., Su, H.-Y., Yan, M.-D., Hsieh, C.-B., Liu, C.-Y., … Lin, Y.-W. (2015). Methylation of IRAK3 is a novel prognostic marker in hepatocellular carcinoma. *World Journal of Gastroenterology : WJG*, *21*(13), 3960–3969. https://doi.org/10.3748/wjg.v21.i13.3960

Kye, B.-H., & Cho, H.-M. (2014). Overview of Radiation Therapy for Treating Rectal Cancer. *Annals of Coloproctology*, *30*(4), 165–174. https://doi.org/10.3393/ac.2014.30.4.165

Lam, K., Pan, K., Linnekamp, J. F., Medema, J. P., & Kandimalla, R. (2016). DNA methylation based biomarkers in colorectal cancer: A systematic review. *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer*, *1866*(1), 106–120. https://doi.org/https://doi.org/10.1016/j.bbcan.2016.07.001

Lambiv, W. L., Vassallo, I., Delorenzi, M., Shay, T., Diserens, A. C., & Misra, A. (2011). The Wnt inhibitory factor 1 (WIF1) is targeted in glioblastoma and has a tumor suppressing function potentially by induction of senescence. *Neuro Oncol*, *13*. https://doi.org/10.1093/neuonc/nor036

Lange, C. P. E., Campan, M., Hinoue, T., Schmitz, R. F., van der Meulen-de Jong, A. E., Slingerland, H., … Laird, P. W. (2012). Genome-Scale Discovery of DNA-Methylation Biomarkers for Blood-Based Detection of Colorectal Cancer. *PLoS ONE*, *7*(11). https://doi.org/10.1371/journal.pone.0050266

Lao, V. V., & Grady, W. M. (2011). Epigenetics and Colorectal Cancer. *Nature Reviews. Gastroenterology & Hepatology*, *8*(12), 686–700. https://doi.org/10.1038/nrgastro.2011.173

Lee, B. B., Lee, E. J., Jung, E. H., Chun, H. K., Chang, D. K., & Song, S. Y. (2009). Aberrant methylation of APC, MGMT, RASSF2A, and Wif-1 genes in plasma as a biomarker for early detection of colorectal cancer. *Clin Cancer Res*, *15*. https://doi.org/10.1158/1078-0432.CCR-09-0111

Lee, B., Godfrey, M., Vitale, E., Hori, H., Mattei, M. G., Sarfarazi, M., … Hollister, D. W. (1991). Linkage of Marfan syndrome and a phenotypically related disorder to two different fibrillin genes. *Nature*, *352*. https://doi.org/10.1038/352330a0

Lee, C. W. L., Chuang, L. S. H., Kimura, S., Lai, S. K., Ong, C. W., Yan, B., … Ito, Y. (2011). RUNX3 functions as an oncogene in ovarian cancer. *Gynecologic Oncology*, *122*(2), 410–417. https://doi.org/http://dx.doi.org/10.1016/j.ygyno.2011.04.044

Lee, G. H., Payne, S. J., Melville, A., & Clark, S. K. (2014). Genetic testing in inherited polyposis syndromes – how and why? *Colorectal Disease*, *16*(8), 595–602. https://doi.org/10.1111/codi.12600

Lengauer, C., Kinzler, K. W., & Vogelstein, B. (1997). Genetic instability in colorectal cancers. *Nature*, *386*(6625), 623–627. Retrieved from http://dx.doi.org/10.1038/386623a0

Lenhard, K., Bommer, G. T., Asutay, S., Schauer, R., Brabletz, T., Göke, B., … Kolligs, F. T. (2005). Analysis of promoter methylation in stool: a novel method for the detection of colorectal cancer. *Clinical Gastroenterology and Hepatology : The Official Clinical Practice Journal of the American Gastroenterological Association*, *3*(2), 142–149. Retrieved from papers://5aecfcca-9729-4def-92fe-c46e5cd7cc81/Paper/p99976

Leong, K. J., Wei, W., Tannahill, L. A., Caldwell, G. M., Jones, C. E., Morton, D. G., … Bach, S. P. (2011). Methylation profiling of rectal cancer identifies novel markers of early-stage disease. *British Journal of Surgery*, *98*(5), 724–734. https://doi.org/10.1002/bjs.7422

Levin, B., Lieberman, D. A., McFarland, B., Andrews, K. S., Brooks, D., Bond, J., … Winawer, S. J. (2008). Screening and surveillance for the early detection of colorectal cancer and adenomatous polyps, 2008: a joint guideline from the American Cancer Society, the US Multi-Society Task Force on Colorectal Cancer, and the American College of Radiology. *Gastroenterology*, *134*. https://doi.org/10.1053/j.gastro.2008.02.002

Li, F., & Lai, M. (2009). Colorectal cancer, one entity or three. *Journal of Zhejiang University SCIENCE B*, *10*(3), 219–229. https://doi.org/10.1631/jzus.B0820273

Li, H., Myeroff, L., Smiraglia, D., Romero, M. F., Pretlow, T. P., Kasturi, L., … Markowitz, S. D. (2003). SLC5A8, a sodium transporter, is a tumor suppressor gene silenced by methylation in human colon aberrant crypt foci and cancers. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(14), 8412–8417. https://doi.org/10.1073/pnas.1430846100

Li, Y.-F., Hsiao, Y.-H., Lai, Y.-H., Chen, Y.-C., Chen, Y.-J., Chou, J.-L., … Tai, C.-K. (2015). DNA methylation profiles and biomarkers of oral squamous cell carcinoma. *Epigenetics*, *10*(3), 229–236. https://doi.org/10.1080/15592294.2015.1006506

Lin, P.-C., Lin, J.-K., Lin, C.-H., Lin, H.-H., Yang, S.-H., Jiang, J.-K., … Chang, S.-C. (2015). Clinical Relevance of Plasma DNA Methylation in Colorectal

Cancer Patients Identified by Using a Genome-Wide High-Resolution Array. *Annals of Surgical Oncology*, *22*(3), 1419–1427. https://doi.org/10.1245/s10434-014-4277-2

Lind, G. E., Ahlquist, T., Kolberg, M., Berg, M., Eknaes, M., Alonso, M. A., … Lothe, R. A. (2008). Hypermethylated MAL gene - a silent marker of early colon tumorigenesis. *J Transl Med*, *6*. https://doi.org/10.1186/1479-5876-6-13

Lind, G. E., Ahlquist, T., & Lothe, R. A. (2007). DNA hypermethylation of MAL: a promising diagnostic biomarker for colorectal tumors. *Gastroenterology*, *132*. https://doi.org/10.1053/j.gastro.2007.03.003

Lind, G. E., Danielsen, S. A., Ahlquist, T., Merok, M. A., Rognum, T. O., Meling, G. I., … Lothe, R. A. (2009). A novel epigenetic biomarker panel for early detection of colorectal cancer and adenomas. *Ejc Supplements*, *7*. https://doi.org/10.1016/S1359-6349(09)70494-X

Lind, G. E., Kleivi, K., Meling, G. I., Teixeira, M. R., Thiis-Evensen, E., Rognum, T. O., & Lothe, R. A. (2006). ADAMTS1, CRABP1, and NR3C1 identified as epigenetically deregulated genes in colorectal tumorigenesis. *Cell Oncol*, *28*.

Lind, G. E., Raiborg, C., Danielsen, S. A., Rognum, T. O., Thiis-Evensen, E., Hoff, G., … Lothe, R. A. (2011). SPG20, a novel biomarker in colorectal carcinogenesis, encodes a regulator of cytokinesis BT - Oncogene.

Lind, G. E., Danielsen, S. A., Ahlquist, T., Merok, M. A., Andresen, K., Skotheim, R. I., … Lothe, R. A. (2011). Identification of an epigenetic biomarker panel with high sensitivity and specificity for colorectal cancer and adenomas. *Molecular Cancer*, *10*(1), 85. https://doi.org/10.1186/1476-4598-10-85

Liu, A. J., Furusato, B., Ravindranath, L., Chen, Y. M., Srikantan, V., McLeod, D. G., … Srivastava, S. (2007). Quantitative analysis of a panel of gene expression in prostate cancer with emphasis on NPY expression analysis. *J Zhejiang Univ Sci B*, *8*. https://doi.org/10.1631/jzus.2007.B0853

Liu, J.-F., Jamieson, G., Wu, T.-C., Zhang, S.-W., Wang, Q.-Z., & Drew, P. (2006). Cyclooxygenase-2 expression in squamous cell carcinoma of the esophagus. *Diseases of the Esophagus*, *19*(5), 350–354. https://doi.org/10.1111/j.1442-2050.2006.00594.x

Livingstone, C. (2013). IGF2 and cancer. *Endocrine-Related Cancer*, *20*(6). https://doi.org/10.1530/ERC-13-0231

Lofton-Day, C., Model, F., DeVos, T., Tetzner, R., Distler, J., Schuster, M., … Sledziewski, A. (2008). DNA methylation biomarkers for blood-based colorectal cancer screening. *Clinical Chemistry*, *54*(2), 414–423. https://doi.org/10.1373/clinchem.2007.095992

Logan, C. Y., & Nusse, R. (2004). The Wnt signaling pathway in development and disease. *Annu Rev Cell Dev Biol*, *20*. https://doi.org/10.1146/annurev.cellbio.20.010403.113126

Lothe, R. A., Peltomäki, P., Meling, G. I., Aaltonen, L. A., Nyström-Lahti, M., Pylkkänen, L., … Børresen, A. L. (1993). Genomic instability in colorectal cancer: relationship to clinicopathological variables and family history. *Cancer Res*, *53*.

Luger, K., Mäder, A. W., Richmond, R. K., Sargent, D. F., & Richmond, T. J. (1997). Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature*, *389*(6648), 251–60. https://doi.org/10.1038/38444

Lunnon, K., Hannon, E., Smith, R. G., Dempster, E., Wong, C., Burrage, J., … Mill, J. (2016). Variation in 5-hydroxymethylcytosine across human cortex

and cerebellum. *Genome Biology*, *17*(1), 27. https://doi.org/10.1186/s13059-016-0871-x
Macrae, F. A., & Bendell, J. (2013). Clinical presentation, diagnosis, and staging of colorectal cancer. (T. K. (Ed), Ed.), *In: UpToDate,*. Retrieved from http://www.uptodate.com/contents/clinical-presentation-diagnosis-and-staging-of-colorectal-cancer?detectedLanguage=en&source=search_result&search=colorectal+cancer&selectedTitle=1~150&provider=noProvider
Mahadevan, V. (2017). The anatomy of the rectum and anal canal. *Surgery - Oxford International Edition*, *29*(1), 5–10. https://doi.org/10.1016/j.mpsur.2010.11.001
Mahmoud, N., Paulson, E., & O'Hara, M. (2017). Colorectal Cancer. Retrieved February 1, 2017, from http://bestpractice.bmj.com/best-practice/monograph/258/diagnosis/step-by-step.html
Mandel, J. S., Bond, J. H., Church, T. R., Snover, D. C., Bradley, G. M., & Schuman, L. M. (2005). Reducing mortality from colorectal cancer by screening for fecal occult blood: Minnesota colon cancer control study. *N Engl J Med*, *328*. https://doi.org/10.1056/NEJM199305133281901
Marie Vedeld, H., Andresen, K., Andrassy Eilertsen, I., Nesbakken, A., Seruca, R., Gladhaug, I. P., … Lind, G. E. (2015). The novel colorectal cancer biomarkers CDO1, ZSCAN18 and ZNF331 are frequently methylated across gastrointestinal cancers. *International Journal of Cancer. Journal International Du Cancer*, *136*(4), 844–853. https://doi.org/10.1002/ijc.29039
Mármol, I., Sánchez-de-Diego, C., Dieste, A. P., Cerrada, E., & Yoldi, M. R. (2017). Colorectal Carcinoma: A General Overview and Future Perspectives in Colorectal Cancer. *International Journal of Molecular Sciences 2017, Vol. 18, Page 197*, *18*(1), 197. https://doi.org/10.3390/IJMS18010197
Matsubara, N. (2012). Epigenetic regulation and colorectal cancer. *Diseases of the Colon and Rectum*, *55*(1), 96–104. https://doi.org/10.1097/DCR.0b013e318233a1ef
Matsuzaki, K., Deng, G., Tanaka, H., Kakar, S., Miura, S., & Kim, Y. S. (2005). The relationship between global methylation level, loss of heterozygosity, and microsatellite instability in sporadic colorectal cancer. *Clinical Cancer Research*, *11*(24), 8564–8569. https://doi.org/10.1158/1078-0432.CCR-05-0859
Maunakea, A. K., Nagarajan, R. P., Bilenky, M., Ballinger, T. J., D'Souza, C., Fouse, S. D., … Costello, J. F. (2010). Conserved Role of Intragenic DNA Methylation in Regulating Alternative Promoters. *Nature*, *466*(7303), 253–257. https://doi.org/10.1038/nature09165
Mazieres, J., He, B., You, L., Xu, Z., Lee, A. Y., & Mikami, I. (2004). Wnt inhibitory factor-1 is silenced by promoter hypermethylation in human lung cancer. *Cancer Res*, *64*. https://doi.org/10.1158/0008-5472.CAN-04-1389
Mc Tavish, N., Copeland, L. A., Saville, M. K., Perkins, N. D., & Spruce, B. A. (2007). Proenkephalin assists stress-activated apoptosis through transcriptional repression of NF-KappaB and p53-related genes targets. *Cell Death Differ*, *14*. https://doi.org/10.1038/sj.cdd.4402172
McLellan, E. A., Medline, A., & Bird, R. P. (1991). Dose response and proliferative characteristics of aberrant crypt foci: putative preneoplastic lesions in rat colon. *Carcinogenesis*, *12*. https://doi.org/10.1093/carcin/12.11.2093

Medical Research Council. (1999). MRC_Human Tissue and Biological Samples for use in Research.
Meling, G. I., Lothe, R. A., Børresen, A. L., Hauge, S., Graue, C., Clausen, O. P., & Rognum, T. O. (1991). Genetic alterations within the retinoblastoma locus in colorectal carcinomas. Relation to DNA ploidy pattern studied by flow cytometric analysis. *Br J Cancer*, *64*. https://doi.org/10.1038/bjc.1991.334
Melki, J. R., Vincent, P. C., & Clark, S. J. (1999). Concurrent DNA hypermethylation of multiple genes in acute myeloid leukemia. *Cancer Res*, *59*.
Mercer, T. R., Dinger, M. E., & Mattick, J. S. (2009). Long non-coding RNAs: insights into functions. *Nature Reviews Genetics*, *10*(3), 155–159. https://doi.org/10.1038/nrg2521
MERCURY Study Group. (2006). Diagnostic accuracy of preoperative magnetic resonance imaging in predicting curative resection of rectal cancer: prospective observational study. *Bmj*, *333*(7572), 779–0. https://doi.org/10.1136/bmj.38937.646400.55
Miao, L., Wang, Y., Xia, H., Yao, C., Cai, H., & Song, Y. (2013). SPOCK1 is a novel transforming growth factor-β target gene that regulates lung cancer cell epithelial-mesenchymal transition. *Biochemical and Biophysical Research Communications*, *440*(4), 792–797. https://doi.org/https://doi.org/10.1016/j.bbrc.2013.10.024
Migliore, L., Migheli, F., Spisni, R., & Copped, F. (2011). Genetics, cytogenetics, and epigenetics of colorectal cancer. *Journal of Biomedicine and Biotechnology*. https://doi.org/10.1155/2011/792362
Mikeska, T., Bock, C., Do, H., & Dobrovic, A. (2012). DNA methylation biomarkers in cancer: progress towards clinical implementation. Retrieved from http://www.expert-reviews.com/doi/abs/10.1586/erm.12.45
Mitchell, S. M., Ross, J. P., Drew, H. R., Ho, T., Brown, G. S., Saunders, N. F. W., … Molloy, P. L. (2014). A panel of genes methylated with high frequency in colorectal cancer. *BMC Cancer*, *14*, 54. https://doi.org/10.1186/1471-2407-14-54
Miyazono, K., Maeda, S., & Imamura, T. (2004). Coordinate regulation of cell growth and differentiation by TGF-beta superfamily and Runx proteins. *Oncogene*, *23*(24), 4232–4237. https://doi.org/10.1038/sj.onc.1207131
Miyazono, K., Maeda, S., & Imamura, T. (n.d.). Coordinate regulation of cell growth and differentiation by TGF-[beta] superfamily and Runx proteins. *Oncogene*, *23*(24), 4232–4237. Retrieved from http://dx.doi.org/10.1038/sj.onc.1207131
Mohammad, H. P., Zhang, W., Prevas, H. S., Leadem, B. R., Zhang, M., Herman, J. G., … Baylin, S. B. (2011). Loss of a single Hic1 allele accelerates polyp formation in Apc(Δ716) mice. *Oncogene*, *30*(23), 2659–2669. https://doi.org/10.1038/onc.2010.633
Mohiuddin, M., & Kramer, S. (1978). Adjuvant radiotherapy--preoperative, postoperative, or both: a proposal for a new approach. *Cancer Clinical Trials*, *1*(2), 93–7.
Molinari, C., Casadio, V., Foca, F., Zingaretti, C., Giannini, M., Avanzolini, A., … Zoli, W. (2013). Gene methylation in rectal cancer: Predictive marker of response to chemoradiotherapy? *Journal of Cellular Physiology*, *228*(12), 2343–2349. https://doi.org/10.1002/jcp.24405
Moon, J. W., Lee, S. K., Lee, J. O., Kim, N., Lee, Y. W., Kim, S. J., … Park, S.-H. (2014). Identification of novel hypermethylated genes and demethylating

effect of vincristine in colorectal cancer. *Journal of Experimental & Clinical Cancer Research : CR*, *33*(1), 4. https://doi.org/10.1186/1756-9966-33-4
Morison, I. M., Ramsay, J. P., & Spencer, H. G. (2005). A census of mammalian imprinting. *Trends in Genetics*, *21*(8), 457–465. https://doi.org/10.1016/j.tig.2005.06.008
Morley-Bunker, A., Walker, L. C., Currie, M. J., Pearson, J., & Eglinton, T. (2016). Translating colorectal cancer genetics into clinically useful biomarkers. *Colorectal Disease*, *18*(8), 749–762. https://doi.org/10.1111/codi.13334
Morris, Z. S., Saha, S., Magnuson, W. J., Morris, B. A., Borkenhagen, J. F., Ching, A., … Ritter, M. A. (2016). Increased tumor response to neoadjuvant therapy among rectal cancer patients taking angiotensin-converting enzyme inhibitors or angiotensin receptor blockers. *Cancer*, *122*(16), 2487–2495. https://doi.org/10.1002/cncr.30079
Müller, H. M., Oberwalder, M., Fiegl, H., Morandell, M., Goebel, G., Zitt, M., … Widschwendter, M. (2004). Methylation changes in faecal DNA: a marker for colorectal cancer screening? *Lancet*, *363*. https://doi.org/10.1016/S0140-6736(04)16002-9
Muller, H. M., & Widschwendter, M. (2003). Methylated DNA as a possible screening marker for neoplastic disease in several body fluids. *Expert Rev Mol Diagn*, *3*. https://doi.org/10.1586/14737159.3.4.443
Mulroney S and Myers A, *Netter's essential physiology E-Book*, Elsevier Health Sciences: 2015.
Mutch, M. G., Schoen, R. E., Fleshman, J. W., Rall, C. J., Dry, S., Seligson, D., … Pinsky, P. F. (2009). A multicenter study of prevalence and risk factors for aberrant crypt foci. *Clin Gastroenterol Hepatol*, *7*. https://doi.org/10.1016/j.cgh.2009.01.016
Nagasaka, T., Tanaka, N., Cullings, H. M., Sun, D.-S., Sasamoto, H., Uchida, T., … Goel, A. (2009). Analysis of Fecal DNA Methylation to Detect Gastrointestinal Neoplasia. *JNCI Journal of the National Cancer Institute*, *101*(18), 1244–1258. https://doi.org/10.1093/jnci/djp265
National Institutes of Health (2005), *Epigenetic mechanisms.* Retrieved from http://commonfund.nih.gov/epigenomics/figure.aspx
Naumov, V. A., Generozov, E. V, Zaharjevskaya, N. B., Matushkina, D. S., Larin, A. K., Chernyshov, S. V, … Govorun, V. M. (2013). Genome-scale analysis of DNA methylation in colorectal cancer using Infinium HumanMethylation450 BeadChips. *Epigenetics*, *8*(9), 921–934. https://doi.org/10.4161/epi.25577
Naviglio, S., Caraglia, M., Abbruzzese, A., Chiosi, E., Di Gesto, D., Marra, M., … Illiano, G. (2009). Protein kinase A as a biological target in cancer therapy. *Expert Opinion on Therapeutic Targets*, *13*(1), 83–92. https://doi.org/10.1517/14728220802602349
Naylor, S. (2003). Biomarkers: current perspectives and future prospects. *Expert Review of Molecular Diagnostics*, *3*(5), 525–529. https://doi.org/10.1586/14737159.3.5.525
Ng, C.-C., Arakawa, H., Fukuda, S., Kondoh, H., & Nakamura, Y. (2003). p53RFP, a p53-inducible RING-finger protein, regulates the stability of p21WAF1. *Oncogene*, *22*(28), 4449–4458. Retrieved from http://dx.doi.org/10.1038/sj.onc.1206586
NICE National Institute for Health Care and Excellence. (2014). *Colorectal cancer: diagnosis and management | Guidance and guidelines | NICE*.

Clinical Guideline (CG131). Retrieved from https://www.nice.org.uk/guidance/cg131
Nojadeh, J. N., Sharif, S. B., & Sakhinia, E. (2018). Microsatellite instability in colorectal cancer. *EXCLI J,* 17, 159-168.
Oberwalder, M., Zitt, M., Wöntner, C., Fiegl, H., Goebel, G., Zitt, M., … Müller, H. M. (2007). SFRP2 methylation in fecal DNA-a marker for colorectal polyps BT - Int J Colorectal Dis.
Ogasawara, M., Murata, J., Ayukawa, K., & Saiki, I. (1997). Differential effect of intestinal neuropeptides on invasion and migration of colon carcinoma cells in vitro. *Cancer Lett*, *119*. https://doi.org/10.1016/S0304-3835(97)81762-4
Ogino, S., Kawasaki, T., Kirkner, G. J., Kraft, P., Loda, M., & Fuchs, C. S. (2007). Evaluation of Markers for CpG Island Methylator Phenotype (CIMP) in Colorectal Cancer by a Large Population-Based Sample. *The Journal of Molecular Diagnostics : JMD*, *9*(3), 305–314. https://doi.org/10.2353/jmoldx.2007.060170
Ogino, S., Kawasaki, T., Kirkner, G. J., Suemoto, Y., Meyerhardt, J. A., & Fuchs, C. S. (2007). Molecular correlates with MGMT promoter methylation and silencing support CpG island methylator phenotype-low (CIMP-low) in colorectal cancer. *Gut*, *56*(11), 1564–71. https://doi.org/10.1136/gut.2007.119750
Oh, T., Kim, N., Moon, Y., Kim, M. S., Hoehn, B. D., Park, C. H., … An, S. (2013). Genome-Wide Identification and Validation of a Novel Methylation Biomarker, SDC2, for Blood-Based Detection of Colorectal Cancer. *The Journal of Molecular Diagnostics*, *15*(4), 498–507. https://doi.org/https://doi.org/10.1016/j.jmoldx.2013.03.004
Oh, T., Kim, N., Moon, Y., Kim, M. S., Hoehn, B. D., Park, C. H., … An, S. (2013). Genome-wide identification and validation of a novel methylation biomarker, SDC2, for blood-based detection of colorectal cancer. *Journal of Molecular Diagnostics*, *15*(4), 498–507. https://doi.org/10.1016/j.jmoldx.2013.03.004
Osborn, N. K., & Ahlquist, D. A. (2005). Stool screening for colorectal cancer: molecular approaches. *Gastroenterology*, *128*. https://doi.org/10.1053/j.gastro.2004.10.041
Øster, B., Thorsen, K., Lamy, P., Wojdacz, T. K., Hansen, L. L., Birkenkamp-Demtrõder, K., … Andersen, C. L. (2011). Identification and validation of highly frequent CpG island hypermethylation in colorectal adenomas and carcinomas BT - Int J Cancer.
Özdemir, B. C., Hensel, J., Secondini, C., Wetterwald, A., Schwaninger, R., Fleischmann, A., … Cecchini, M. G. (2014). The Molecular Signature of the Stroma Response in Prostate Cancer-Induced Osteoblastic Bone Metastasis Highlights Expansion of Hematopoietic and Prostate Epithelial Stem Cell Niches. *PLoS ONE*, *9*(12), e114530. https://doi.org/10.1371/journal.pone.0114530
Pangeni, R. P., Channathodiyil, P., Huen, D. S., Eagles, L. W., Johal, B. K., Pasha, D., … Morris, M. R. (2015). The GALNT9, BNC1 and CCDC8 genes are frequently epigenetically dysregulated in breast tumours that metastasise to the brain. *Clinical Epigenetics*, *7*(1), 57. https://doi.org/10.1186/s13148-015-0089-x
Park, S. Y., Chang, H. J., Kim, D. Y., Jung, K. H., Kim, S. Y., Park, J. W., … Jeong, S. Y. (2011). Is step section necessary for determination of complete pathological response in rectal cancer patients treated with

preoperative chemoradiotherapy? *Histopathology*, *59*(4), 650–659. https://doi.org/10.1111/j.1365-2559.2011.03980.x
Parkin, D. M., Bray, F., Ferlay, J., & Pisani, P. (2001). Estimating the world cancer burden:Globocan 2000. *Int J Cancer*, *94*. https://doi.org/10.1002/ijc.1440
Patel, U. B., Taylor, F., Blomqvist, L., George, C., Evans, H., Tekkis, P., … Brown, G. (2011). Magnetic resonance imaging-detected tumor response for locally advanced rectal cancer predicts survival outcomes: MERCURY experience. *Journal of Clinical Oncology*, *29*(28), 3753–3760. https://doi.org/10.1200/JCO.2011.34.9068
Paul, C. L., & Clark, S. J. (1996). Cytosine methylation: quantitation by automated genomic sequencing and GENESCAN analysis. *BioTechniques*, *21*(1), 126–33. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8816247
Pavicic, W., Joensuu, E. I., Nieminen, T., & Peltomaki, P. (2012). LINE-1 hypomethylation in familial and sporadic cancer. *J Mol Med (Berl)*, *90*. https://doi.org/10.1007/s00109-011-0854-z
Payne, S. R. (2010). From discovery to the clinic: the novel DNA methylation biomarker (m)SEPT9 for the detection of colorectal cancer in blood. *Epigenomics*, *2*(4), 575–85. https://doi.org/10.2217/epi.10.35
Pedersen, B. S., Schwartz, D. A., Yang, I. V., & Kechris, K. J. (2012). Comb-p: Software for combining, analyzing, grouping and correcting spatially correlated P-values. *Bioinformatics*, *28*(22), 2986–2988. https://doi.org/10.1093/bioinformatics/bts545
Pedersen, S. K., Symonds, E. L., Baker, R. T., Murray, D. H., McEvoy, A., Van Doorn, S. C., … Young, G. P. (2015). Evaluation of an assay for methylated BCAT1 and IKZF1 in plasma for detection of colorectal neoplasia. *BMC Cancer*, *15*(1), 654. https://doi.org/10.1186/s12885-015-1674-2
Perea, J., Cano, J. M., Rueda, D., García, J. L., Inglada, L., Osorio, I., … Urioste, M. (2015). Classifying early-onset colorectal cancer according to tumor location: new potential subcategories to explore. *American Journal of Cancer Research*, *5*(7), 2308–2313. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4548343/
Petersen, S. H., Harling, H., Kirkeby, L. T., Wille-Jørgensen, P., & Mocellin, S. (2012). Postoperative adjuvant chemotherapy in rectal cancer operated for cure. *Cochrane Database of Systematic Reviews*, (3). https://doi.org/10.1002/14651858.CD004078.pub2
Pidsley, R., Y Wong, C. C., Volta, M., Lunnon, K., Mill, J., & Schalkwyk, L. C. (2013). A data-driven approach to preprocessing Illumina 450K methylation array data. *BMC Genomics*, *14*(1), 293. https://doi.org/10.1186/1471-2164-14-293
Polakis, P. (2012). Wnt signaling in cancer. *Cold Spring Harb Perspect Biol*, *4:a008052p*(4:a008052), 4:a008052. https://doi.org/10.1038/onc.2016.304
Popović, A., Demirović, A., Spajić, B., Štimac, G., Krušlin, B., & Tomas, D. (2010). Expression and prognostic role of syndecan-2 in prostate cancer. *Prostate Cancer and Prostatic Diseases*, *13*(1), 78–82. https://doi.org/10.1038/pcan.2009.43
Poston, G. J., Tait, D., O'Connell, S., Bennett, A., & Berendse, S. (2011). Diagnosis and management of colorectal cancer: summary of NICE guidance. *BMJ*, *343*. Retrieved from http://www.bmj.com/content/343/bmj.d6751.abstract

Powell, S. M. (1992). APC mutations occur early during colorectal tumorigenesis. *Nature*, *359*. https://doi.org/10.1038/359235a0
Prasanth, K. V., & Spector, D. L. (2007). Eukaryotic regulatory RNAs: An answer to the “genome complexity” conundrum. *Genes and Development*, *21*(1), 11–42. https://doi.org/10.1101/gad.1484207
Pretlow, T. P., & Pretlow, T. G. (2005). Mutant KRAS in aberrant crypt foci (ACF): initiation of colorectal cancer? *Biochim Biophys Acta*, *1756*.
Qi, J., Zhu, Y. Q., Luo, J., & Tao, W. H. (2006). Hypermethylation and expression regulation of secreted frizzled-related protein genes in colorectal tumor. *World Journal of Gastroenterology*, *12*(44), 7113–7117. https://doi.org/10.3748/wjg.v12.i44.7113
QIAamp. (2012). QIAamp DNA FFPE tissue handbook.
QIAGEN. (2012). QIAGEN Allprep DNA/RNA mini kit handbook.
Qin, Q., Xu, Y., He, T., Qin, C., & Xu, J. (2012). Normal and disease-related biological functions of Twist1 and underlying molecular mechanisms. *Cell Research*, *22*(1), 90–106. https://doi.org/10.1038/cr.2011.144
Qin, Y., Zhang, Q., Lee, S., Zhong, W., & Liu, Y. (2015). Doxycycline reverses epithelial-to-mesenchymal transition and suppresses the proliferation and metastasis of lung cancer cells. *Oncotarget*, *6*(38).
Quintanilla, I., Lopez-Cerón, M., Jimeno, M., Cuatrecasas, M., Muñoz, J., Moreira, L., … Balaguer, F. (2014). LINE-1 hypomethylation is neither present in rectal aberrant crypt foci nor associated with field defect in sporadic colorectal neoplasia. *Clinical Epigenetics*, *6*(1), 24. https://doi.org/10.1186/1868-7083-6-24
Ramírez, N., Bandrés, E., Navarro, A., Pons, A., Jansa, S., Moreno, I., … García-Foncillas, J. (2008). Epigenetic events in normal colonic mucosa surrounding colorectal cancer lesions. *European Journal of Cancer*, *44*(17), 2689–2695. https://doi.org/10.1016/j.ejca.2008.09.004
Rao, C. N., Cook, B., Liu, Y., Chilukuri, K., Stack, M. S., Foster, D. C., … Woodley, D. T. (1998). HT-1080 fibrosarcoma cell matrix degradation and invasion are inhibited by the matrix-associated serine protease inhibitor TFPI-2/33 kDa MSPI. *International Journal of Cancer*, *76*(5), 749–756. https://doi.org/10.1002/(SICI)1097-0215(19980529)76:5<749::AID-IJC21>3.0.CO;2-Y
Rao, C. N., Reddy, P., Reeder, D. J., Liu, Y., Stack, S. M., Kisiel, W., & Woodley, D. T. (2000). Prokaryotic Expression, Purification, and Reconstitution of Biological Activities (Antiprotease, Antitumor, and Heparin-Binding) for Tissue Factor Pathway Inhibitor-2. *Biochemical and Biophysical Research Communications*, *276*(3), 1286–1294. https://doi.org/10.1006/bbrc.2000.3611
Rashid, A., Shen, L., Morris, J. S., Issa, J. P., & Hamilton, S. R. (2001). CpG island methylation in colorectal adenomas. *Am.J.Pathol.*, *159*(0002–9440 SB–AIM SB–IM), 1129–1135.
Rauscher, G. H., Kresovich, J. K., Poulin, M., Yan, L., Macias, V., Mahmoud, A. M., … Ehrlich, M. (2015). Exploring DNA methylation changes in promoter, intragenic, and intergenic regions as early and late events in breast cancer formation. *BMC Cancer*, *15*(1), 816. https://doi.org/10.1186/s12885-015-1777-9
Rawson, J. B., Manno, M., Mrkonjic, M., Daftary, D., Dicks, E., Buchanan, D. D., … Bapat, B. (2011). Promoter methylation of Wnt antagonists DKK1 and SFRP1 is associated with opposing tumor subtypes in two large

populations of colorectal cancer patients. *Carcinogenesis*, *32*(5), 741–747. https://doi.org/10.1093/carcin/bgr020
Razin, A., & Cedar, H. (1991). DNA methylation and gene expression. *Microbiological Reviews*, *55*(3), 451–458. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC372829/
Renehan, A. G., Tyson, M., Egger, M., Heller, R. F., & Zwahlen, M. (2008). Body-mass index and incidence of cancer: a systematic review and meta-analysis of prospective observational studies. *Lancet*, *371*. https://doi.org/10.1016/S0140-6736(08)60269-X
Renehan, A. G., Malcomson, L., Emsley, R., Gollins, S., Maw, A., Myint, A. S., … O'Dwyer, S. T. (n.d.). Watch-and-wait approach versus surgical resection after chemoradiotherapy for patients with rectal cancer (the OnCoRe project): a propensity-score matched cohort analysis. *The Lancet Oncology*. https://doi.org/10.1016/S1470-2045(15)00467-2
Reya, T., & Clevers, H. (2005). Wnt signalling in stem cells and cancer. *Nature*, *434*(7035), 843–850.
Rhim, A. D., Oberstein, P. E., Thomas, D. H., Mirek, E. T., Palermo, C. F., Sastra, S. A., … Stanger, B. Z. (2014). Stromal elements act to restrain, rather than support, pancreatic ductal adenocarcinoma. *Cancer Cell*, *25*(6), 735–747. https://doi.org/10.1016/j.ccr.2014.04.021
Richards, E. J. (2006). Inherited epigenetic variation--revisiting soft inheritance. *Mycorrhiza*, *7*(5), 395–401. https://doi.org/10.1038/nrg1834
Rodriguez-Bigas, M., & Grothey, A. (2017). Overview of the management of primary colon cancer. In D. Savarese (Ed.). Waltham MA: UpTodate.
Rogers, M. A., Kalter, V., Strowitzki, M., Schneider, M., & Lichter, P. (2016). IGF2 knockdown in two colorectal cancer cell lines decreases survival, adhesion and modulates survival-associated genes. *Tumor Biology*, *37*(9), 12485–12495. https://doi.org/10.1007/s13277-016-5115-x
ROMANOSKI, C. E., GLASS, C. K., STUNNENBERG, H. G., WILSON, L., & ALMOUZNI, G. (2015). Epigenomics: Roadmap for regulation. *Nature*, *518*, 314–6.
Roperch, J.-P., Incitti, R., Forbin, S., Bard, F., Mansour, H., Mesli, F., … Sobhani, I. (2013). Aberrant methylation of NPY, PENK, and WIF1 as a promising marker for blood-based diagnosis of colorectal cancer. *BMC Cancer*, *13*(1), 566. https://doi.org/10.1186/1471-2407-13-566
Sakurai, K., Urade, M., Noguchi, K., Hashitani, S., Takaoka, K., Segawa, E., & Kishimoto, H. (n.d.). Prognostic significance of cyclooxygenase-2 and DNA topoisomerase IIalpha expression in oral carcinoma. *Head Neck*, *29*(11), 1002–1009.
Salerno, G., Sinnatamby, C., Branagan, G., Daniels, I. R., Heald, R. J., & Moran, B. J. (2006). Defining the rectum: surgically, radiologically and anatomically. *Colorectal Disease*, *8*, 5–9. https://doi.org/10.1111/j.1463-1318.2006.01062.x
Samad, A. K. A., Taylor, R. S., Marshall, T., & Chapman, M. A. S. (2005). A meta-analysis of the association of physical activity with reduced risk of colorectal cancer. *Colorectal Disease*, *7*(3), 204–213. https://doi.org/10.1111/j.1463-1318.2005.00747.x
Sansom, O. J. (2004). Loss of Apc in vivo immediately perturbs Wnt signaling, differentiation, and migration. *Genes Dev*, *18*. https://doi.org/10.1101/gad.287404
Sasaki, H., Jones, P. A., Chaillet, J. R., Ferguson-Smith, A. C., Barton, S. C., Reik, W., & Surani, M. A. (1992). Parental imprinting: Potentially active

chromatin of the repressed maternal allele of the mouse insulin-like growth factor II (Igf2) gene. *Genes and Development*, *6*(10), 1843–1856. https://doi.org/10.1101/gad.6.10.1843

Sauer, R., Becker, H., Hohenberger, W., Rödel, C., Wittekind, C., Fietkau, R., … Raab, R. (2004). Preoperative versus Postoperative Chemoradiotherapy for Rectal Cancer. *New England Journal of Medicine*, *351*(17), 1731–1740. https://doi.org/10.1056/NEJMoa040694

Schrag, D. (2013). Evolving Role of Neoadjuvant Therapy in Rectal Cancer. *Current Treatment Options in Oncology*, *14*(3), 350–364. https://doi.org/10.1007/s11864-013-0242-8

Schwarzenbach, H., Hoon, D. S. B., & Pantel, K. (2011). Cell-free nucleic acids as biomarkers in cancer patients. *Nat Rev Cancer*, *11*(6), 426–437. Retrieved from http://dx.doi.org/10.1038/nrc3066

Sebio, a, Salazar, J., Páez, D., Berenguer-Llergo, A., Del Río, E., Tobeña, M., … Barnadas, A. (2014). EGFR ligands and DNA repair genes: genomic predictors of complete response after capecitabine-based chemoradiotherapy in locally advanced rectal cancer. *The Pharmacogenomics Journal*, (January), 1–7. https://doi.org/10.1038/tpj.2014.33

Segditsas, S., & Tomlinson, I. (2006). Colorectal cancer and genetic alterations in the Wnt pathway. *Oncogene*, *25*. https://doi.org/10.1038/sj.onc.1210059

Shames, D. S., Girard, L., Gao, B., Sato, M., Lewis, C. M., Shivapurkar, N., … Minna, J. D. (2006). A genome-wide screen for promoter methylation in lung cancer identifies novel methylation markers for multiple malignancies. *PLoS Med*, *3*. https://doi.org/10.1371/journal.pmed.0030486

Shan, T., Cui, X., Li, W., Lin, W., & Li, Y. (2014). AQP5: a novel biomarker that predicts poor clinical outcome in colorectal cancer. *Oncol Rep*, *32*.

Sharma, S., Kelly, T. K., & Jones, P. A. (2010). Epigenetics in cancer. *Carcinogenesis*, *31*(1), 27–36. https://doi.org/10.1093/carcin/bgp220

Sheaffer, K. L., Elliott, E. N., & Kaestner, K. H. (2016). DNA hypomethylation contributes to genomic instability and intestinal cancer initiation. *Cancer Prevention Research*, *9*(7), 534–546. https://doi.org/10.1158/1940-6207.CAPR-15-0349

Shin, S. H., Kim, B. H., Jang, J. J., Suh, K. S., & Kang, G. H. (2010). Identification of novel methylation markers in hepatocellular carcinoma using a methylation array. *J Korean Med Sci*, *25*. https://doi.org/10.3346/jkms.2010.25.8.1152

Shoemaker, R., Deng, J., Wang, W., & Zhang, K. (2010). Allele-specific methylation is prevalent and is contributed by CpG-SNPs in the human genome. *Genome Research*, *20*(7), 883–889. https://doi.org/10.1101/gr.104695.109

Shukla, S., Kavak, E., Gregory, M., Imashimizu, M., Shutinoski, B., Kashlev, M., … Oberdoerffer, S. (2011). CTCF-promoted RNA polymerase II pausing links DNA methylation to splicing. *Nature*, *479*(7371), 74–79. Retrieved from http://dx.doi.org/10.1038/nature10442

Sidak, Z. (1967). Rectangular confidence region for the means of multivariate normal distributions. *Journal of the American Statistical Association*, *62*, 626–633.

Siegel, R. L., Miller, K. D., & Jemal, A. (2017). Cancer statistics, 2017. *CA: A Cancer Journal for Clinicians*, *67*(1), 7–30. https://doi.org/10.3322/caac.21387

Simmer, F., Brinkman, A. B., Assenov, Y., Matarese, F., Kaan, A., Sabatino, L., … Stunnenberg, H. G. (2012). Comparative genome-wide DNA methylation analysis of colorectal tumor and matched normal tissues, (December), 1355–1367.

Sizer, B. F., Arulampalam, T., Austin, R., Lacey, N., Menzies, D., & Motson, R. (2006). MRI in predicting curative resection of rectal cancer: defining a "window of opportunity" for laparoscopic surgery. *BMJ (Clinical Research Ed.)*, *333*(7572), 808–809. https://doi.org/10.1136/bmj.333.7572.808-b

Slattery, M. L., Lundgreen, A., Herrick, J. S., Caan, B. J., Potter, J. D., & Wolff, R. K. (2011). Associations between genetic variation in RUNX1, RUNX2, RUNX3, MAPK1 and eIF4E and risk of colon and rectal cancer: additional support for a TGF-β-signaling pathway. *Carcinogenesis*, *32*(3), 318–326. https://doi.org/10.1093/carcin/bgq245

Slaughter, D. P., Southwick, H. W., & Smejkal, W. (1953). ???Field cancerization??? in oral stratified squamous epithelium. Clinical implications of multicentric origin. *Cancer*, *6*(5), 963–968. https://doi.org/10.1002/1097-0142(195309)6:5<963::AID-CNCR2820060515>3.0.CO;2-Q

Smith, N., & Brown, G. (2008). Preoperative staging of rectal cancer. *Acta Oncologica (Stockholm, Sweden)*, *47*(1), 20–31. https://doi.org/10.1080/02841860701697720

Smith, R. A., Andrews, K. S., Brooks, D., Fedewa, S. A., Manassaram-Baptiste, D., Saslow, D., … Wender, R. C. (2017). Cancer screening in the United States, 2017: A review of current American Cancer Society guidelines and current issues in cancer screening. *CA: A Cancer Journal for Clinicians*, *67*(2), 100–121. https://doi.org/10.3322/caac.21392

Sobhani, I., Alzahouri, K., Ghout, I., Charles, D. J., & Durand-Zaleski, I. (2011). Cost-effectiveness of mass screening for colorectal cancer: choice of fecal occult blood test and screening strategy. *Dis Colon Rectum*, *54*. https://doi.org/10.1007/DCR.0b013e31820fd2bc

Sonneberg, J. L., Rauscher, F. J., Morgan, J. I., & Curran, T. (1989). Regulation of proenkephalin by fos and Jun. *Science*, *246*. https://doi.org/10.1126/science.2512642

Sparmann, A., & van Lohuizen, M. (2006). Polycomb silencers control cell fate, development and cancer. *Nature Reviews Cancer*, *6*(11), 846–856. https://doi.org/10.1038/nrc1991

Speek, M. (2001). Antisense promoter of human L1 retrotransposon drives transcription of adjacent cellular genes. *Mol Cell Biol*, *21*. https://doi.org/10.1128/MCB.21.6.1973-1985.2001

Spencer, V. A., & Davie, J. R. (1999). Role of covalent modifications of histones in regulating gene expression. *Gene*, *240*(1), 1–12. https://doi.org/10.1016/S0378-1119(99)00405-9

Srinivasan, M., Sedmak, D., & Jewell, S. (2002). Effect of Fixatives and Tissue Processing on the Content and Integrity of Nucleic Acids. *The American Journal of Pathology*, *161*(6), 1961–1971. https://doi.org/10.1016/S0002-9440(10)64472-0

Stankevicius, V., Vasauskas, G., Noreikiene, R., Kuodyte, K., Valius, M., & Suziedelis, K. (2016). Extracellular matrix-dependent pathways in colorectal cancer cell lines reveal potential targets for anticancer therapies. *Anticancer Research*, *36*(9), 4559–4567. https://doi.org/10.21873/anticanres.11004

Steemers, F. J., Chang, W., Lee, G., Barker, D. L., Shen, R., & Gunderson, K. L. (2006). Whole-genome genotyping with the single-base extension assay. *Nature Methods*, *3*(1), 31–33. https://doi.org/10.1038/nmeth842

Strahl, B. D., & Allis, C. D. (2000). The language of covalent histone modifications. *Nature*, *403*(6765), 41–5. https://doi.org/10.1038/47412

Strimbu, K., & Tavel, J. A. (2010). What are Biomarkers? *Current Opinion in HIV and AIDS*, *5*(6), 463–466. https://doi.org/10.1097/COH.0b013e32833ed177

Subramaniam, M. M., Chan, J. Y., Soong, R., Ito, K., Yeoh, K. G., Wong, R., … Salto-Tellez, M. (2009). RUNX3 Inactivation in Colorectal Polyps Arising Through Different Pathways of Colonic Carcinogenesis. *The American Journal of Gastroenterology*, *104*(2), 426–436. https://doi.org/10.1038/ajg.2008.141

Subramaniam, M. M., Chan, J. Y., Yeoh, K. G., Quek, T., Ito, K., & Salto-Tellez, M. (2009). Molecular pathology of RUNX3 in human carcinogenesis. *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer*, *1796*(2), 315–331. https://doi.org/https://doi.org/10.1016/j.bbcan.2009.07.004

Suehiro, Y., & Hinoda, Y. (2008). Genetic and epigenetic changes in aberrant crypt foci and serrated polyps. *Cancer Sci*, *99*. https://doi.org/10.1111/j.1349-7006.2008.00784.x

Sugai, T., Habano, W., Takagi, R., Yamano, H., Eizuka, M., Arakawa, N., … Matsumoto, T. (2017). Analysis of molecular alterations in laterally spreading tumors of the colorectum. *Journal of Gastroenterology*, *52*(6), 715–723. https://doi.org/10.1007/s00535-016-1269-y

Sun, W., Sun, Y., Zhu, M., Wang, Z., Zhang, H., Xin, Y., … Liu, Y. (2014). The role of plasma cell-free DNA detection in predicting preoperative chemoradiotherapy response in rectal cancer patients. *Oncol Rep*, *31*(3), 1466–1472. https://doi.org/10.3892/or.2013.2949

Sunami, E., de Maat, M., Vu, A., Turner, R. R., & Hoon, D. S. (2011). LINE-1 hypomethylation during primary colon cancer progression. *PLoS One*, *6*. https://doi.org/10.1371/journal.pone.0018884

Suzuki, H., Gabrielson, E., Chen, W., Anbazhagan, R., van Engeland, M., Weijenberg, M. P., … Baylin, S. B. (2002). A genomic screen for genes upregulated by demethylation and histone deacetylase inhibition in human colorectal cancer. *Nat Genet*, *31*. https://doi.org/10.1038/ng892

Suzuki, H., Watkins, D. N., Jair, K.-W., Schuebel, K. E., Markowitz, S. D., Dong Chen, W., … Baylin, S. B. (2004). Epigenetic inactivation of SFRP genes allows constitutive WNT signaling in colorectal cancer. *Nature Genetics*, *36*(4), 417–422. https://doi.org/10.1038/ng1330

Suzuki, M. M., Bird, A., Hu, F., Ge, S., Ye, M., Xiang, H., … Kumar, S. (2008). DNA methylation landscapes: provocative insights from epigenomics. *Nature Reviews Genetics*, *9*(6), 465–476. https://doi.org/10.1038/nrg2341

Szmulowicz, U. M., & Hull, T. L. (2011). Colonic Physiology. In D. E. Beck, P. L. Roberts, T. J. Saclarides, A. J. Senagore, M. J. Stamos, & S. D. Wexner (Eds.), *The ASCRS Textbook of Colon and Rectal Surgery: Second Edition* (pp. 23–39). New York, NY: Springer New York. https://doi.org/10.1007/978-1-4419-1584-9_2

Takayama, T., Katsuki, S., Takahashi, Y., Ohi, M., Nojiri, S., Sakamaki, S., … Niitsu, Y. (1998). Aberrant crypt foci of the colon as precursors of adenoma and cancer. *N Engl J Med*, *339*. https://doi.org/10.1056/NEJM199810293391803

Tamas, K., Walenkamp, A. M. E., de Vries, E. G. E., van Vugt, M. A. T. M., Beets-Tan, R. G., van Etten, B., … Hospers, G. A. P. (2017). Rectal and colon cancer: Not just a different anatomic site. *Cancer Treatment Reviews*, *41*(8), 671–679. https://doi.org/10.1016/j.ctrv.2015.06.007

Tan, A. C., Jimeno, A., Lin, S. H., Wheelhouse, J., Chan, F., & Solomon, A. (2009). Characterizing DNA methylation patterns in pancreatic cancer genome. *Mol Oncol*, *3*. https://doi.org/10.1016/j.molonc.2009.03.004

Tan, S. H., Ida, H., Lau, Q. C., Goh, B. C., Chieng, W. S., Loh, M., & Ito, Y. (2007). Detection of promoter hypermethylation in serum samples of cancer patients by methylation-specific polymerase chain reaction for tumour suppressor genes including RUNX3. *Oncology Reports*, *18*(5), 1225–1230.

Tang, D., Liu, J., Wang, D. R., Yu, H. F., Li, Y. K., & Zhang, J. Q. (2011). Diagnostic and prognostic value of the methylation status of secreted frizzled-related protein 2 in colorectal cancer. *Clin Invest Med*, *34*.

Taniguchi, H., Yamamoto, H., Hirata, T., Miyamoto, N., Oki, M., & Nosho, K. (2005). Frequent epigenetic inactivation of Wnt inhibitory factor-1 in human gastrointestinal cancers. *Oncogene*, *24*. https://doi.org/10.1038/sj.onc.1208910

Tänzer, M., Balluff, B., Distler, J., Hale, K., Leodolter, A., Röcken, C., … Ebert, M. P. A. (2010). Performance of epigenetic markers SEPT9 and ALX4 in plasma for detection of colorectal precancerous lesions. *PLoS ONE*, *5*(2). https://doi.org/10.1371/journal.pone.0009061

Tardivo, S., Mantovani, W., Torri, E., & Poli, A. (2005). Rectal Cancer. Epidemiology and Burden of Disease BT - Rectal Cancer: New Frontiers in Diagnosis, Treatment and Rehabilitation. In G. G. Delaini (Ed.) (pp. 1–21). Milano: Springer Milan. https://doi.org/10.1007/88-470-0343-1_1

Taylor, F., George, C., Guthrie, A., Bees, N., Ub, P., Lk, B., & Taylor, F. (2012). MRI After Treatment of Locally Advanced Rectal Cancer: How to Report Tumor Response— The MERCURY Experience. *Ajr American Journal Of Roentgenology*, (October), 1. https://doi.org/10.2214/AJR.11.8210

Tegeder, I., & Geisslinger, G. (2004). Opioids as modulators of cell death and survival-unraveling mechanisms and revealing new indications. *Pharmacol Rev*, *56*. https://doi.org/10.1124/pr.56.3.2

Thiis-Evensen, E., Hoff, G. S., Sauar, J., Langmark, F., Majak, B. M., & Vatn, M. H. (1999). Population-based surveillance by colonoscopy: effect on the incidence of colorectal cancer. Telemark Polyp Study I. *Scand J Gastroenterol*, *34*. https://doi.org/10.1080/003655299750026443

Thorstensen, L., Lind, G. E., Løvig, T., Diep, C. B., Meling, G. I., Rognum, T. O., & Lothe, R. A. (2005). Genetic and Epigenetic Changes of Components Affecting the WNT Pathway in Colorectal Carcinomas Stratified by Microsatellite Instability. *Neoplasia*, *7*. https://doi.org/10.1593/neo.04448

Thrumurthy, S. G., Thrumurthy, S. S. D., Gilbert, C. E., Ross, P., & Haji, A. (2016). Colorectal adenocarcinoma: risks, prevention and diagnosis. *BMJ*, *354*. Retrieved from http://www.bmj.com/content/354/bmj.i3590.abstract

Toyota, M., Ahuja, N., Ohe-Toyota, M., Herman, J. G., Baylin, S. B., & Issa, J.-P. J. (1999). CpG island methylator phenotype in colorectal cancer. *Proceedings of the National Academy of Sciences of the United States of America*, *96*(15), 8681–8686. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC17576/

Toyota, M., Ho, C., Ahuja, N., Jair, K. W., Li, Q., Ohe-Toyota, M., … Issa, J. P. J. (1999). Identification of differentially methylated sequences in colorectal

cancer by methylated CpG island amplification. *Cancer Research*, *59*(10), 2307–2312.
Toyota, M., & Issa, J.-P. J. (1999). CpG island methylator phenotypes in aging and cancer. *Seminars in Cancer Biology*, *9*(5), 349–357. https://doi.org/10.1006/scbi.1999.0135
Tsang, J. S., Vencken, S., Sharaf, O., Leen, E., Kay, E. W., McNamara, D. A., … Mulligan, E. D. (2017). Global DNA methylation is altered by neoadjuvant chemoradiotherapy in rectal cancer and may predict response to treatment – A pilot study. *European Journal of Surgical Oncology*, *40*(11), 1459–1466. https://doi.org/10.1016/j.ejso.2014.06.008
Tsikitis, V. L., Malireddy, K., Green, E. A., Christensen, B., Whelan, R., Hyder, J., … Nelson, H. (2009). Postoperative Surveillance Recommendations for Early Stage Colon Cancer Based on Results From the Clinical Outcomes of Surgical Therapy Trial. *Journal of Clinical Oncology*, *27*(22), 3671–3676. https://doi.org/10.1200/JCO.2008.20.7050
Tsuchiya, N., Ochiai, M., Nakashima, K., Ubagai, T., Sugimura, T., & Nakagama, H. (2007). SND1, a component of RNA-induced silencing complex, is up-regulated in human colon cancers and implicated in early stage colon carcinogenesis. *Cancer Research*, *67*(19), 9568–9576. https://doi.org/10.1158/0008-5472.CAN-06-2707
Tsunematsu, T., Kudo, Y., Iizuka, S., Ogawa, I., Fujita, T., Kurihara, H., … Takata, T. (2009). RUNX3 has an oncogenic role in head and neck cancer. *PLoS ONE*, *4*(6). https://doi.org/10.1371/journal.pone.0005892
Ueki, T., Toyata, M., & Skinner, H. (2001). Identification and characterization of differentially methylated CpG islands in pancreatic carcinoma. *Cancer Res*, *61*.
Uen, Y.-H., Fang, C.-L., Hseu, Y.-C., Shen, P.-C., Yang, H.-L., Wen, K.-S., … Lin, K.-Y. (2015). VAV3 Oncogene Expression in Colorectal Cancer: Clinical Aspects and Functional Characterization. *Scientific Reports*, *5*, 9360. https://doi.org/10.1038/srep09360
Uhl, B., Dietrich, D., Branchi, V., Semaan, A., Schaefer, P., Gevensleben, H., … Matthaei, H. (2016). DNA Methylation of PITX2 and PANCR Is Prognostic for Overall Survival in Patients with Resected Adenocarcinomas of the Biliary Tract. *PLoS ONE*, *11*(10), e0165769. https://doi.org/10.1371/journal.pone.0165769
Ukkola, O., Ravussin, E., Jacobson, P., Perusse, L., Rankinen, T., & Tschop, M. (2002). Role of ghrelin polymorphisms in obesity based on three different studies. *Obes Res*, *10*. https://doi.org/10.1038/oby.2002.106
University, S. L. (2017). Department of Biochemsitry & Molecular Biology Research Programs.
van, D. L., Kuipers, E. J., & van Leerdam, M. E. (2010). Performance improvements of stool-based screening tests. *Best Pract Res Clin Gastroenterol*, *24*.
Veeck, J., Wild, P. J., Fuchs, T., Schuffler, P. J., Hartmann, A., Knüchel, R., & Dahl, E. (2009). Prognostic relevance of Wnt-inhibitory factor-1 (WIF1) and Dickkopf-3 (DKK3) promoter methylation in human breast cancer. *BMC Cancer*, *9*. https://doi.org/10.1186/1471-2407-9-217
Virani, S., Colacino, J. A., Kim, J. H., & Rozek, L. S. (2012). Cancer Epigenetics: A Brief Review. *ILAR Journal*, *53*(3–4), 359–369. https://doi.org/10.1093/ilar.53.3-4.359
Vymetalkova, V., Vodicka, P., Pardini, B., Rosa, F., Levy, M., Schneiderova, M., … Farkas, S. A. (2016). Epigenome-wide analysis of DNA methylation

reveals a rectal cancer-specific epigenomic signature. *Epigenomics*, *8*(9), 1193–1207. https://doi.org/10.2217/epi-2016-0044
Waddington, C. H. (1957). *The Strategy of the Genes: A Discussion of Some Aspects of Theoretical Biology.* London: Allen & Unwin.
Waddington, C. H. (1942). The epigenotype. *Endeavour*, *1*, 18–20.
Wagman, R., Minsky, B. D., Cohen, A. M., Guillem, J. G., & Paty, P. P. (1998). Sphincter preservation in rectal cancer with preoperative radiation therapy and coloanal anastomosis: long term follow-up. *International Journal of Radiation Oncology • Biology • Physics*, *42*(1), 51–57. https://doi.org/10.1016/S0360-3016(98)00180-1
Wahab, A. H. A., El-mezayen, H. A., Sharad, H., & Rahman, S. A. (2011). Promoter hypermethylation of RASSF1A, MGMT, and HIC-1 genes in benign and malignant colorectal tumors. *Tumor Biology*, *32*(5), 845–852. https://doi.org/10.1007/s13277-011-0156-7
Wang, D. R., & Tang, D. (2008). Hypermethylated SFRP2 gene in fecal DNA is a high potential biomarker for colorectal cancer noninvasive screening. *World J Gastroenterol*, *14*. https://doi.org/10.3748/wjg.14.524
Wang, R. Y. H., Gehrke, C. W., & Ehrlich, M. (1980). Comparison of bisulfite modification of 5-methyldeoxycytidine and deoxycytidine residues. *Nucleic Acids Research*, *8*(20), 4777–4790. https://doi.org/10.1093/nar/8.20.4777
Wang, X. W., Wei, W., Wang, W. Q., Zhao, X. Y., Guo, H., & Fang, D. C. (2014). RING finger proteins are involved in the progression of Barrett esophagus to esophageal adenocarcinoma: A preliminary study. *Gut and Liver*, *8*(5), 487–494. https://doi.org/10.5009/gnl13133
Wang, Y., Yu, Q., Cho, A. H., Rondeau, G., Welsh, J., Adamson, E., … McClelland, M. (2005). Survey of differentially methylated promoters in prostate cancer cell lines. *Neoplasia*, *7*. https://doi.org/10.1593/neo.05289
Warren, J. D., Xiong, W., Bunker, A. M., Vaughn, C. P., Furtado, L. V, Roberts, W. L., … Heichman, K. A. (2011). Septin 9 methylated DNA is a sensitive and specific blood test for colorectal cancer. *BMC Medicine*, *9*, 133. https://doi.org/10.1186/1741-7015-9-133
Warwick, R., Williams, P. L., Dyson, M., & Bannister, L. (1985). *Gray's anatomy*. *British Journal of Radiology* (Vol. 58). Retrieved from http://bjr.birjournals.org/cgi/content/abstract/58/694/1029
Watt, F., & Molloy, P. L. (1988). Cytosine methylation prevents binding to DNA of a HeLa cell transcription factor required for optimal expression of the adenovirus major late promoter. *Genes & Development*, *2*(9), 1136–1143. https://doi.org/10.1101/gad.2.9.1136
Wei, E. K., Giovannucci, E., Wu, K., Rosner, B., Fuchs, C. S., Willett, W. C., & Colditz, G. A. (2004). COMPARISON OF RISK FACTORS FOR COLON AND RECTAL CANCER. *International Journal of Cancer. Journal International Du Cancer*, *108*(3), 433–442. https://doi.org/10.1002/ijc.11540
Wei, J., Li, G., Zhang, J., Zhou, Y., Dang, S., Chen, H., … Liu, M. (2016). Integrated analysis of genome-wide DNA methylation and gene expression profiles identifies potential novel biomarkers of rectal cancer. *Oncotarget*, *7*(38), 62547–62558. https://doi.org/10.18632/oncotarget.11534
Weisenberger, D. J., Campan, M., Long, T. I., Kim, M., Woods, C., Fiala, E., … Laird, P. W. (2005). Analysis of repetitive element DNA methylation by MethyLight. *Nucleic Acids Res*, *33*. https://doi.org/10.1093/nar/gki987
Weisenberger, D. J., Siegmund, K. D., Campan, M., Young, J., Long, T. I., Faasse, M. a, … Laird, P. W. (2006). CpG island methylator phenotype underlies sporadic microsatellite instability and is tightly associated with

BRAF mutation in colorectal cancer. *Nature Genetics*, *38*(7), 787–93. https://doi.org/10.1038/ng1834
Wendt, M. K., Johanesen, P. a, Kang-Decker, N., Binion, D. G., Shah, V., & Dwinell, M. B. (2006). Silencing of epithelial CXCL12 expression by DNA hypermethylation promotes colonic carcinoma metastasis. *Oncogene*, *25*(36), 4986–97. https://doi.org/10.1038/sj.onc.1209505
Wheeler, J. M. D. (2005). Epigenetics, mismatch repair genes and colorectal cancer. *Annals of The Royal College of Surgeons of England*, *87*(1), 15–20. https://doi.org/10.1308/1478708051423
Widschwendter, M., Siegmund, K. D., Muller, H. M., Fiegl, H., Marth, C., Muller-Holzner, E., … Laird, P. W. (2004). Association of breast cancer DNA methylation profiles with hormone receptor status and response to tamoxifen. *Cancer Res*, *64*. https://doi.org/10.1158/0008-5472.CAN-03-3852
Williamson, J. S., Harris, D. A., Beynon, J., & Jenkins, G. J. S. (2015). Review of the development of DNA methylation as a marker of response to neoadjuvant therapy and outcomes in rectal cancer. *Clinical Epigenetics*, *7*(1), 70. https://doi.org/10.1186/s13148-015-0111-3
Wilson, A. S., Power, B. E., & Molloy, P. L. (2007). DNA hypomethylation and human diseases. *North*, *1775*, 138–162. https://doi.org/10.1016/j.bbcan.2006.08.007
Winawer, S., Fletcher, R., Rex, D., Bond, J., Burt, R., & Ferrucci, J. (2003). Colorectal cancer screening and surveillance: clinical guidelines and rationale-update based on new evidence. *Gastroenterology*, *124*. https://doi.org/10.1053/gast.2003.50044
Winslow, T. (n.d.). Colorectal Cancer: Tumour sizes.
Wissmann, C., Wild, P. J., Kaiser, S., Roepcke, S., Stoehr, R., & Woenckhaus, M. (2003). WIF1, a component of the Wnt pathway, is down-regulated in prostate, breast, lung, and bladder cancer. *J Pathol*, *201*. https://doi.org/10.1002/path.1449
Wolin, K. Y., Yan, Y., Colditz, G. A., & Lee, I.-M. (2009). Physical activity and colon cancer prevention: a meta-analysis. *British Journal of Cancer*, *100*(4), 611–616. https://doi.org/10.1038/sj.bjc.6604917
Wong, C. M., Ng, Y. L., Lee, J. M. F., Wong, C. C. L., Cheung, O. F., Chan, C. Y., … Ng, I. O. L. (2007). Tissue factor pathway inhibitor-2 as a frequently silenced tumour suppressor gene in hepatocellular carcinoma. *Hepatology*, *45*(5), 1129–1138. https://doi.org/10.1002/hep.21578
Wong, J. J., Hawkins, N. J., & Ward, R. L. (2007). Colorectal cancer: a model for epigenetic tumorigenesis. *Gut*, *56*. https://doi.org/10.1136/gut.2005.088799
Worthley, D. L., Whitehall, V. L. J., Buttenshaw, R. L., Irahara, N., Greco, S. A., Ramsnes, I., … Leggett, B. A. (2010). DNA methylation within the normal colorectal mucosa is associated with pathway-specific predisposition to cancer. *Oncogene*, *29*(11), 1653–1662. https://doi.org/10.1038/onc.2009.449
Wu, B., Piloto, S., Zeng, W., Hoverter, N. P., Schilling, T. F., & Waterman, M. L. (2013). Ring Finger Protein 14 is a new regulator of TCF/β-catenin-mediated transcription and colon cancer cell survival. *EMBO Reports*, *14*(4), 347–355. https://doi.org/10.1038/embor.2013.19
Wu, J. S. (2007). Rectal Cancer Staging. *Clinics in Colon and Rectal Surgery*, *20*(3), 148–157. https://doi.org/10.1055/s-2007-984859

Wu, J., Wang, G., He, B., Chen, X., & An, Y. (2017). Methylation of the UNC5C gene and its protein expression in colorectal cancer. *Tumor Biology*, *39*(4), 101042831769756. https://doi.org/10.1177/1010428317697564

Xie, C., Yuan, J., Li, H., Li, M., Zhao, G., Bu, D., … Zhao, Y. (2014). NONCODEv4: Exploring the world of long non-coding RNA genes. *Nucleic Acids Research*, *42*(D1). https://doi.org/10.1093/nar/gkt1222

Xu, Y., Hu, B., Choi, A. J., Gopalan, B., Lee, B. H., Kalady, M. F., … Ting, A. H. (2012). Unique DNA methylome profiles in CpG island methylator phenotype colon cancers. *Genome Research*, *22*(2), 283–291. https://doi.org/10.1101/gr.122788.111

Yamada, A., Minamiguchi, S., Sakai, Y., Horimatsu, T., Muto, M., Chiba, T., … Goel, A. (2014). Colorectal advanced neoplasms occur through dual carcinogenesis pathways in individuals with coexisting serrated polyps. *PLoS One*, *9*. https://doi.org/10.1371/journal.pone.0098059

Yamashita, K., Waraya, M., Sook Kim, M., Sidransky, D., Katada, N., Sato, T., … Watanabe, M. (2014). Detection of methylated CDO1 in plasma of colorectal cancer; A PCR study. *PLoS ONE*, *9*(12). https://doi.org/10.1371/journal.pone.0113546

Yamauchi, M., Morikawa, T., Kuchiba, A., Imamura, Y., Qian, Z. R., Nishihara, R., … Ogino, S. (2012). Assessment of Colorectal Cancer Molecular Features along Bowel Subsites Challenges the Conception of Distinct Dichotomy of Proximal vs. Distal Colorectum. *Gut*, *61*(6), 847–854. https://doi.org/10.1136/gutjnl-2011-300865

Yan, P. S., Venkataramu, C., Ibrahim, A., Liu, J. C., Shen, R. Z., Diaz, N. M., … Huang, T. H. (2006). Mapping geographic zones of cancer risk with epigenetic biomarkers in normal breast tissue. *Clin Cancer Res*, *12*. https://doi.org/10.1158/1078-0432.CCR-06-0467

Yang, L., Lee, M. M. K., Leung, M. M. H., & Wong, Y. H. (2016). Regulator of G protein signaling 20 enhances cancer cell aggregation, migration, invasion and adhesion. *Cellular Signalling*, *28*(11), 1663–1672. https://doi.org/https://doi.org/10.1016/j.cellsig.2016.07.017

Yang, Z., Wang, Y., Fang, J., Chen, F., Liu, J., & Wu, J. (2010). Expression and aberrant promoter methylation of Wnt inhibitory factor-1 in human astrocytomas. *J Exp Clin Cancer Res*, *29*. https://doi.org/10.1186/1756-9966-29-26

Yeo, S.-G., Kim, D. Y., Kim, T. H., Chang, H. J., Oh, J. H., Park, W., … Lee, K.-J. (2010). Pathologic complete response of primary tumor following preoperative chemoradiotherapy for locally advanced rectal cancer: long-term outcomes and prognostic significance of pathologic nodal status (KROG 09-01). *Annals of Surgery*, *252*(6), 998–1004. https://doi.org/10.1097/SLA.0b013e3181f3f1b1

Yi, J. M., Dhir, M., Van Neste, L., Downing, S. R., Jeschke, J., Glöckner, S. C., … Ahuja, N. (2011). Genomic and epigenomic integration identifies a prognostic signature in colon cancer. *Clinical Cancer Research : An Official Journal of the American Association for Cancer Research*, *17*(6), 1535–1545. https://doi.org/10.1158/1078-0432.CCR-10-2509

Yi, J. M., & Kim, T. O. (2015). Epigenetic Alterations in Inflammatory Bowel Disease and Cancer. *Intestinal Research*, *13*(2), 112–121. https://doi.org/10.5217/ir.2015.13.2.112

Yoder, J. A., Walsh, C. P., & Bestor, T. H. (2017). Cytosine methylation and the ecology of intragenomic parasites. *Trends in Genetics*, *13*(8), 335–340. https://doi.org/10.1016/S0168-9525(97)01181-5

Yokoi, K., Yamashita, K., Ishii, S., Tanaka, T., Nishizawa, N., Tsutsui, A., … Watanabe, M. (2017). Comprehensive molecular exploration identified promoter DNA methylation of the CRBP1 gene as a determinant of radiation sensitivity in rectal cancer. *British Journal of Cancer*. https://doi.org/10.1038/bjc.2017.65

Yokota, T., Sugano, K., Kondo, H., Saito, D., Sugihara, K., Fukayama, N., … Yoshida, S. (1997). Detection of aberrant crypt foci by magnifying colonoscopy. *Gastrointest Endosc*, *46*. https://doi.org/10.1016/S0016-5107(97)70212-8

Yu, F., Li, G., Gao, J., Sun, Y., Liu, P., Gao, H., … Mu, Q. (2016). SPOCK1 is upregulated in recurrent glioblastoma and contributes to metastasis and Temozolomide resistance. *Cell Proliferation*, *49*(2), 195–206. https://doi.org/10.1111/cpr.12241

Zakhari, S. (2013). Epigenetics: New Frontier for Alcohol Research. *Alcohol Research: Current Reviews*, *35*(1), 1–10.

Zampieri, M., Ciccarone, F., Calabrese, R., Franceschi, C., B??rkle, A., & Caiafa, P. (2015). Reconfiguration of DNA methylation in aging. *Mechanisms of Ageing and Development*, *151*, 60–70. https://doi.org/10.1016/j.mad.2015.02.002

Zanella, E. R., Galimi, F., Sassi, F., Migliardi, G., Cottino, F., Leto, S. M., … Bertotti, A. (2015). IGF2 is an actionable target that identifies a distinct subpopulation of colorectal cancer patients with marginal response to anti-EGFR therapies. *Science Translational Medicine*, *7*(272), 272ra12 LP-272ra12. Retrieved from http://stm.sciencemag.org/content/7/272/272ra12.abstract

Zhan, T., Rindtorff, N., & Boutros, M. (2017, March). Wnt signaling in cancer. *Oncogene*. The Author(s). Retrieved from http://dx.doi.org/10.1038/onc.2016.304

Zhang, H., Song, Y. C., & Dang, C. X. (2013). Detection of hypermethylated spastic paraplegia-20 in stool samples of patients with colorectal cancer. *International Journal of Medical Sciences*, *10*(3), 230–234. https://doi.org/10.7150/ijms.5278

Zhang, W., Bauer, M., Croner, R. S., Pelz, J. O. W., Lodygin, D., Hermeking, H., … Matzel, K. E. (2007). DNA stool test for colorectal cancer: Hypermethylation of the secreted frizzled-related protein-1 gene. *Diseases of the Colon and Rectum*, *50*(10), 1618–1626. https://doi.org/10.1007/s10350-007-0286-6

Zhang, X., Song, Y.-F., Lu, H.-N., Wang, D.-P., Zhang, X.-S., Huang, S.-L., … Huang, Z.-G. (2015). Combined detection of plasma GATA5 and SFRP2 methylation is a valid noninvasive biomarker for colorectal cancer and adenomas. *World Journal of Gastroenterology : WJG*, *21*(9), 2629–2637. https://doi.org/10.3748/wjg.v21.i9.2629

Zhong, H., Fazenbaker, C., Chen, C., Breen, S., Huang, J., Yao, X., … Hollingsworth, R. E. (2017, February 9). Overproduction of IGF-2 drives a subset of colorectal cancer cells, which specifically respond to an anti-IGF therapeutic antibody and combination therapies. *Oncogene*. Macmillan Publishers Limited. Retrieved from http://dx.doi.org/10.1038/onc.2016.248

Zhong, H., Fazenbaker, C., Chen, C., Breen, S., Huang, J., Yao, X., … Hollingsworth, R. E. (2017, February 9). Overproduction of IGF-2 drives a subset of colorectal cancer cells, which specifically respond to an anti-IGF therapeutic antibody and combination therapies. *Oncogene*. Macmillan Publishers Limited. Retrieved from http://dx.doi.org/10.1038/onc.2016.248

ZHU, Z., MU, Y., QI, C., WANG, J., XI, G., GUO, J., … ZHAO, F. (2015). CYP1B1 enhances the resistance of epithelial ovarian cancer cells to paclitaxel in vivo and in vitro. *International Journal of Molecular Medicine*, *35*(2), 340–348. https://doi.org/10.3892/ijmm.2014.2041

Ziller, M. J., Gu, H., Müller, F., Donaghey, J., Tsai, L. T.-Y., Kohlbacher, O., … Meissner, A. (2013). Charting a dynamic DNA methylation landscape of the human genome. *Nature*, *500*(7463), 477–81. https://doi.org/10.1038/nature12433

Zitt, M., Zitt, M., & Muller, H. M. (2007). DNA methylation in colorectal cancer--impact on screening and therapy monitoring modalities? *Dis Markers*, *23*. https://doi.org/10.1155/2007/891967

Zou, H., Harrington, J. J., Shire, A. M., Rego, R. L., Wang, L., Campbell, M. E., … Ahlquist, D. A. (2007). Highly methylated genes in colorectal neoplasia: implications for screening. *Cancer Epidemiol Biomarkers Prev*, *16*. https://doi.org/10.1158/1055-9965.EPI-07-0518